PROBLEMS IN ESTIMATION THEORY WITH APPLICATIONS TO ORBIT DETERMINATION

David Walker Curkendall

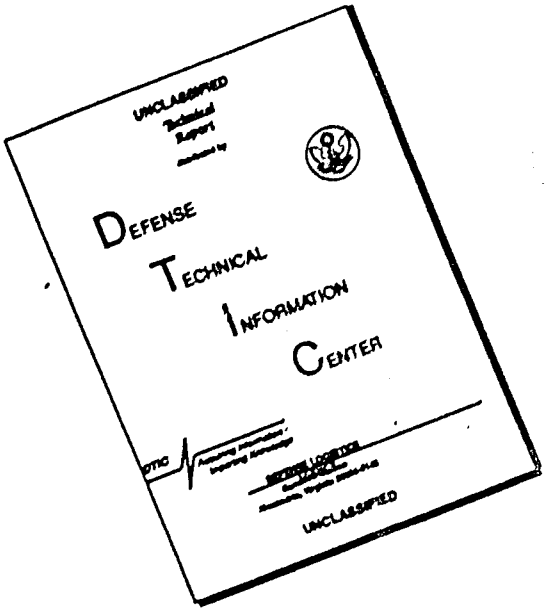
California University

Prepared for:

Air Force Office of Scientific Research
Air Force Space and Missile Systems Organization

September 1972

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

UCLA-ENG-7275
SEPTEMBER 1972

PROBLEMS IN ESTIMATION THEORY WITH APPLICATIONS TO ORBIT DETERMINATION

D.W. CURKENDALL

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)
School of Engineering and Applied Science
University of California
Los Angeles, California 90024

2a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

2b. GROUP

3. REPORT TITLE

PROBLEMS IN ESTIMATION THEORY WITH APPLICATIONS TO ORBIT DETERMINATION

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Scientific Interim

5. AUTHOR(S) (First name, middle initial, last name)
David W. Curkendall

6. REPORT DATE
September 1972

7a. TOTAL NO. OF PAGES
182

7b. NO. OF REFS
26

8a. CONTRACT OR GRANT NO.
AFOSR 72-2166

b. PROJECT NO. 9769

c. 61102F

d. 681304

9a. ORIGINATOR'S REPORT NUMBER(S)
UCLA-ENG 7275

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
AFOSR - TR - 73 - 0268

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES
TECH., OTHER

12. SPONSORING MILITARY ACTIVITY
Air Force Office of Scientific Research/NM
1400 Wilson Boulevard
Arlington, Virginia 22209

13. ABSTRACT

Several problems in estimation theory with current application interest to orbit determination are treated in this dissertation. Although much of the material covered has general application, the important specific of interplanetary orbit determination is given special emphasis and is used throughout to provide context and to obtain numerical results. Topics covered include: modeling the errors in the navigation measurements, describing the navigation measurements compactly in terms of the unknown parameters, and describing the estimation accuracy in a variety of poorly modeled situations.

The phase-coherent doppler system used by NASA's Deep Space Net is modeled as a time-invariant linear system to determine the influence the master oscillator stability has on the resultant doppler data accuracy.

Analytical expressions for determining the errors introduced in spacecraft state estimates due to Earth-based tracking station location errors and constant unmodeled forces acting on the probe are derived. This is accomplished through re-consideration and extension of time explicit analytic expressions for the doppler and range observables as a function of the probe state.

The Householder transformation technique for the solution to the linear least squares problem is recalled. This method has been shown to be numerically superior to either the classical normal equations or standard Kalman filtering approaches to the solution of the least squares problem by avoiding certain squaring operations in the computational algorithm. Algorithms are derived which permit a wider range of applicability of the Householder technique to estimation problems.

DD FORM 1 NOV 68 1473

Ib

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Estimation Theory Kalman Filtering Orbital Determination						

UCLA-ENG-7275
September 1972

PROBLEMS IN ESTIMATION THEORY WITH APPLICATIONS TO ORBIT DETERMINATION

David W. Curkendall

School of Engineering and Applied Science
University of California
Los Angeles

PREFACE

The constant search for the development of efficient and improved estimation techniques for sensor data in U.S. Air Force Systems is an effort which will continue indefinitely with, presumably, ever improving results. This report develops a rather comprehensive set of new results in this area of fundamental importance to so many U.S. Air Force applications including Space Systems, Missile systems, tactical systems, aeronautical systems, and other U.S. Air Force systems although the specific vehicle for application throughout this report is space systems.

This research report was prepared under research projects supported by the U.S. Air Force Office of Scientific Research under AFOSR Grant 72-2166, Design of Aerospace Systems, and by the U.S. Air Force Space and Missile Systems Organization under Contract No. F04701-72-C-0273, Advanced Space Guidance, and this report constitutes part of the final report for these contracts.

The research described in this report "Problems In Estimation Theory With Application To Orbit Determination," UCLA-ENG-7275, by David Walter Curkendall, was carried out under the direction of C.T. Leondes and E.B. Stear, Co-Principal Investigators in the Schools of Engineering in the University of California at Los Angeles and Santa Barbara, respectively.

ABSTRACT

Several problems in estimation theory with current application interest to orbit determination are treated in this dissertation. Although much of the material covered has general application, the important specific of interplanetary orbit determination is given special emphasis and is used throughout to provide context and to obtain numerical results. Topics covered include: modeling the errors in the navigation measurements, describing the navigation measurements compactly in terms of the unknown parameters, and describing the estimation accuracy in a variety of poorly modeled situations.

The phase-coherent doppler system used by NASA's Deep Space Net is modeled as a time-invariant linear system to determine the influence the master oscillator stability has on the resultant doppler data accuracy.

Analytical expressions for determining the errors introduced in spacecraft state estimates due to Earth-based tracking station location errors and constant unmodeled forces acting on the probe are derived. This is accomplished through re-consideration and extension of time explicit analytic expressions for the doppler and range observables as a function of the probe state.

The Householder transformation technique for the solution to the linear least squares problem is recalled. This method has been shown to be numerically superior to either the classical normal equations or standard Kalman filtering approaches to the solution of the least

squares problem by avoiding certain squaring operations in the computational algorithm. Algorithms are derived which permit a wider range of applicability of the Householder technique to estimation problems. Specifically, methods are provided for calculating the actual estimation performance when the apriori modeling used by the filter is incorrect or certain stochastic processes affecting either the data directly or the probe state itself are present. These algorithms are used in several key numerical examples calculating the effects of stochastic station location motion and probe state process noise.

The problem of detecting the presence of parameters not modeled by the filter through the inspection of the data residuals is considered. Again, the Householder transformation technique is used, this time to make the problem of computing the expected sum-of-squares of the residuals as a function of the unmodeled parameters computationally tractable. Several suggestions are made concerning this detection problem and an informal hypothesis testing procedure for validating least-squares solutions is detailed.

CONTENTS

FIGURES

FIGURES (Cont'd)

FIGURES (Cont'd)

TABLES

CHAPTER I

INTRODUCTION

Navigation is an integral part of the design and execution of any space mission. The ability to determine and control the flight path of the spacecraft to high accuracy is a prime mission requirement in order that the mission objectives can be met. In all but the simplest unmanned missions, orbit determination and trajectory control are needed, not just during the initial powered flight phase, but throughout the mission, so that continuing refinements in knowledge or changes to the flight path can be made to optimize the scientific return. For manned missions, the additional concerns for crew safety add to the imperative reasons for these functions.

The current state of the art of navigation systems design result in mission performance limited not by the mechanization errors in the sub-systems that control the changes in the flight path, but by the orbit determination accuracy. The above statement is a generalization and cannot be taken categorically since it is possible to construct sensible mission situations where the reverse would be true. Still in all demanding interplanetary missions currently planned - Mariner/Mars '71, Mariner/Venus/Mercury '73, Viking '75, Grand Tour Missions, etc. - a factor of 10 improvement in the control sub-systems would refine mission performance while a similar factor in orbit determination accuracy would revolutionize it. A more complete description of the state of the art and current inherent

limitations of planetary navigation is found in Reference 1.

Estimation theory as applied to orbit determination is selected as the subject for the dissertation because of this importance to overall mission design and performance. Several key problems relative to earth-based, radio navigation are investigated and new results are obtained from both the estimation theoretic and physical model points of view. A discussion of the content and major results contained in the following chapters is:

Chapter II

The phase-coherent doppler tracking system is modeled as a linear, time-invariant system and the effects that variations in the output frequency of the master oscillator have on the doppler measurement are determined. It is found that the slow drifts in the oscillator frequency are of greater importance than the relatively much larger, but more rapid, frequency excursions arising from the wide-band component of the oscillator's output spectrum.

In general, the results presented permit the determination of the oscillator performance characteristics required to achieve a specified doppler data accuracy.

Chapter III

Series expansion methods are used in order to obtain a time-explicit analytic relationship for the doppler observable in terms of the (unknown) probe-state and the earth-based tracking station coordinates. This is an extension of earlier work by Hamilton and Melbourne (Reference 2). The resulting expansion is used as the basis for an analytic orbit determination accuracy analysis which

reveals the structure of the accuracies to be expected as a function of the probe-Earth-Sun geometry. The analysis is approximate and only valid for relatively short spans of tracking data (less than one week for a probe in interplanetary space) but is highly useful in gaining intuitive understanding, establishing the key error sources, and interpreting results from more precise analyses. The effects of augmenting the basic doppler data type with ranging data are explored.

Chapter IV

The analysis of the previous chapter is employed to derive analytic expressions for the errors in the estimate of probe state as a function of the two most important model errors: 1) station location errors and 2) unmodeled forces affecting the probe's motion. This analysis assumes doppler-only tracking; similar analytic results for the case where ranging data are also employed do not appear tractable. Because the structure of the analytic analysis suggests that ranging data could reduce the error sensitivities to the unmodeled forces, a numerical study is presented to test this hypothesis. The results show that the position errors are reduced by this additional data in the vicinity of the data itself, but at the expense of larger velocity errors. These results lay the foundation for the intelligent application of this supplementary data type in operational situations.

Chapter V

A short history of the development of square-root filtering using Householder transformations is given and the applicability of

this more numerically stable computational approach is extended to include two important problems. They are: to evaluate true filter performance when 1) the apriori uncertainty assigned to the estimated variable in filter initialization is modeled improperly and 2) Markovian, stochastic processes, of which the filter is unaware, are affecting the data and/or the state. These extensions, when combined with algorithms currently available, provide the basis for the investigation of the succeeding chapter.

Chapter VI

The limitations of the analytic analyses of Chapters III and IV permit computations involving neither long spans of tracking data nor the evaluation of the effects of stochastic processes. To circumvent this, a more precise, but non-analytic, model is adopted to investigate the effects of stochastically varying station location errors and unmodeled process noise as a function of the length of the data arc. It is found that:

1. Increasing the length of the data arc decreases the errors due to both fixed and random station location errors in two of the three components of position. The error in the third coordinate (probe declination) is remarkably insensitive to data span and shows little variation over the entire range investigated (up to 4 months tracking).
2. Solving for the fixed components of station location errors produces better performance even in the presence of the random components which are ignored by

the filter. But again, the improvement is selective (e.g., a factor of 2.3 in probe right ascension but only a factor of 1.3 in declination).

3. The sensitivity of estimation error to constant, unmodeled forces increases by a factor of four from that predicted by the short-term analytic analysis after approximately three weeks of data are processed. However, qualitative structure of the analytic result in terms of which components of the solution are affected is found to be accurate even for relatively long arcs of data.

4. For a given magnitude of the acceleration affecting the probe, unmodeled stochastically varying accelerations produce approximately a factor of from ten to forty greater estimation errors than do unmodeled constant accelerations. However, by including ranging data and extending the filter to include solutions for constant accelerations, the effects of the stochastic forces can be reduced dramatically (by a factor of ten in the case studied.)

Because the above results are not obtained with an analytic analysis, it is difficult to determine their universality in different mission applications with different geometries. However, it is thought that the techniques developed are applicable to a wide range of interplanetary orbit determination problems and can be a significant aid to forging appropriate estimation strategies.

Chapter VII

Householder transformation techniques are used to determine the behavior of the minimum weighted sum of squares of the data residuals in the presence of unmodeled parameters. In addition, formulae are derived which yield the decrease in the sum-of-squares which would be obtained by including additional parameters in the solution without actually performing the extended solutions. These results are employed in a suggested informal hypothesis testing procedure whose aims are to arrive at a reliable data set in the sense that inspection of the residual behavior will reveal modeling difficulties and to conduct the search for the source of the difficulty efficiently.

Appendix

Much of the dissertation assumes a general familiarity with the earth-based data types being analyzed and estimation methods used in orbit determination. This background material is reviewed in the Appendix. Both range and doppler data are discussed and a brief presentation of how the orbit determination problem is structured for computational solution using digital computers is given. Simple derivations of classical weighted least squares and sequential filtering are presented and these methods are related to the square-root filtering algorithms employed throughout the dissertation.

CHAPTER II.

THE INFLUENCE OF MASTER OSCILLATOR STABILITY ON DOPPLER DATA ACCURACY

The main data type obtained during Earth-based spacecraft tracking is the two-way coherent doppler signal which is proportional to the relative range rate between the station and the spacecraft. In actual practice, this doppler tone is integrated by means of an electronic counter, producing the raw data type which is proportional to the range change between these two participants over the counting interval. A significant contributor to the total doppler system error is the frequency instability in the station master oscillator which supplies the frequency reference for both the original transmission to the spacecraft and for comparison with the return signal to detect the doppler tone.

The objective of this chapter is to apply linear system analysis techniques to the doppler system in order to determine the quantitative relationship between the statistical description of the oscillator performance and the resultant integrated doppler data errors. Both two-way coherent tracking (just described) and three-way tracking (where the receiving station is at a distant location and by necessity employs an independent reference oscillator for the doppler detection) are investigated. The case of white noise in the oscillator output is treated quite thoroughly, but it is shown that the resulting data noise, while quite visible in the data is not as serious as that

arising from an oscillator error output of primarily low frequency content.*

Linear System Analysis

From the doppler system diagram shown and discussed in the Appendix, the two-way, coherent tracking system can be modeled as shown in Figure 2.1. Here, the error signal, $f_e(t)$ is fed into a delay line of duration ℓ (the round trip light time to the probe) and simultaneously fed toward to be differenced with the output of the delay line. The result is the instantaneous doppler error, $\dot{\rho}_e$, which in turn is integrated to produce a range-change, or integrated doppler error, $\Delta\rho_e$. The lower diagram (ii) is the same system and constitutes the definitions of $h_1(t)$ and $h_2(t)$, the unit impulse response of the two subsystems. From inspection:

$$h_1(t) = \delta(t) - \delta(t - \ell) \qquad (2.1)$$

Assume the application of a mean zero wide sense-(or covariance-) stationary process of spectral composition, $G_f(\omega)$(the mean zero assumption is not really necessary since it is readily seen that the steady-state output of the system to a constant input is zero). Recalling the relation

$$G_{\dot{\rho}}(\omega) = G_f(\omega)\,|H_1(j\omega)|^2$$

*We have to contend here with the rather confusing semantic problem that the error output is a frequency. In modeling this error as a stochastic process we frequently must refer to its power spectral density and speak of its frequency content. This dual use of the term frequency is unfortunate but apparently unavoidable. We shall attempt however, to speak particularly carefully in this regard to keep the confusion at a minimum.

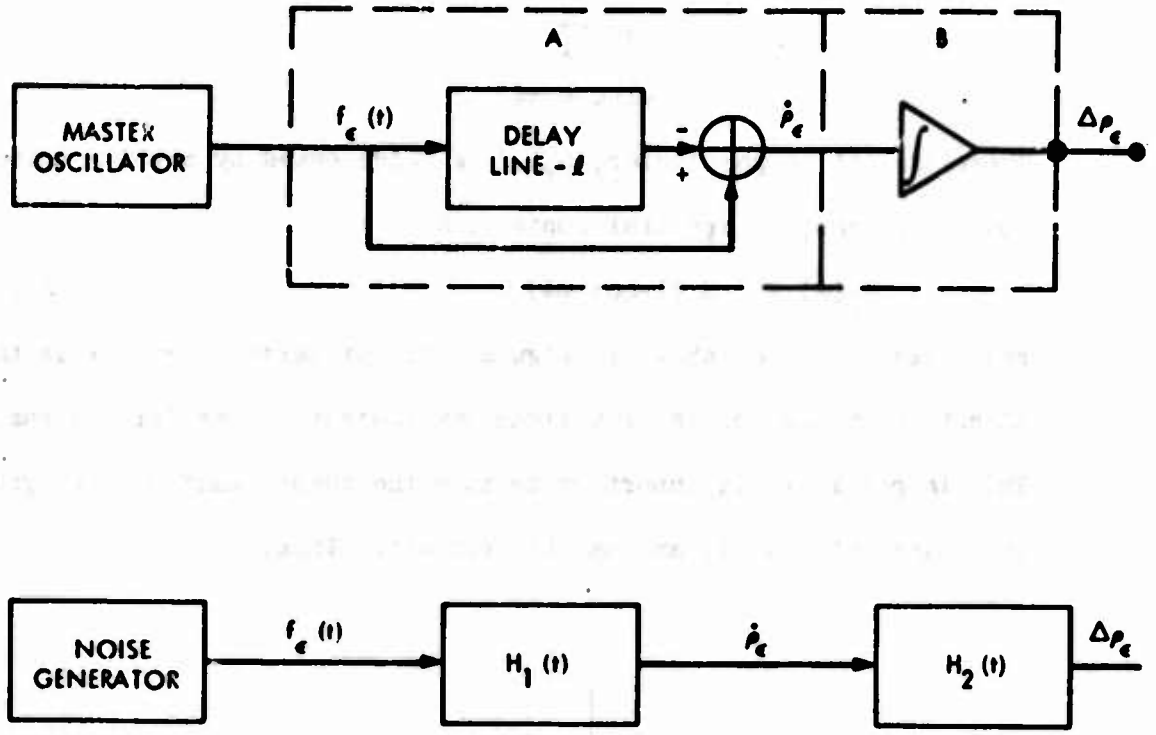


Figure 2.1. Block Diagram of Doppler System.

where $H_1(j\omega)$ is the system function and is defined as the Fourier transform of $h_1(t)$.

$$H_1(j\omega) = \int_{-\infty}^{\infty} \left[\delta(t) - \delta(t-\ell)\right] e^{-j\omega t} dt$$

$$= 1 - e^{-j\omega\ell}$$

$$|H_1(j\omega)|^2 = H_1(j\omega) H_1^{*}(j\omega)$$

$$= 2\,(1-\cos\omega\ell$$

Therefore, if in particular, $G_{\dot{\rho}}(\omega)$ is represented by a white noise source of constant spectral content, K

$$G_{\dot{\rho}}(\omega) = 2\,K\,(1-\cos\omega\ell) \qquad (2.2)$$

and would appear as shown in Figure 2.2. Of particular note is the attentuation made of the low frequency content of the input signal. This is particularly important because the second half of the system is a pure integrator, an unstable element. Thus,

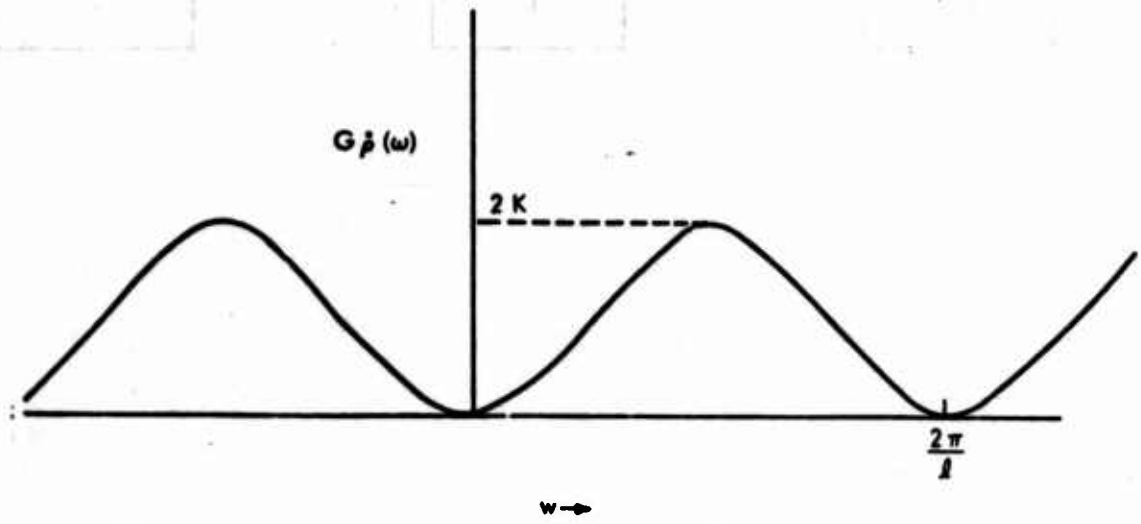


Figure 2.2. Instantaneous Doppler Signal Error From a White Noise Oscillator Error Model.

suppose that $G_\rho(\omega)$ contained a component of k'/ω^2 (real oscillators are supposed to contain a component $\simeq \frac{1}{\omega}$, the so-called "flicker-noise" component). Since

$$1-\cos \omega l \approx \frac{\omega^2 l^2}{2}$$

the spectral height of $G_\rho(o)$ would still be finite. This observation is in the nature of an aside and may appear unmotivated, but it helps to demonstrate the important point that bias errors in the doppler data are extremely low in magnitude. To demonstrate their existence at all, second order arguments must be pursued.

Continuing with the white noise postulate, it is useful to take the inverse Fourier transform of (2) to obtain the autocorrelation function of the doppler, i.e.

$$\begin{aligned} R_\rho(\tau) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} G_\rho(\omega) e^{j\omega\tau} d\omega \\ &= \frac{K}{2\pi}\int_{-\infty}^{\infty} (2-e^{-j\omega l}-e^{j\omega l}) e^{j\omega\tau} d\omega \\ &= \frac{K}{\pi}\left[\delta(\tau) - \frac{1}{2}\delta(\tau - l) - \frac{1}{2}\delta(\tau + l)\right] \end{aligned} \qquad (2.3)$$

This function is normalized and displayed in Figure 2.3.

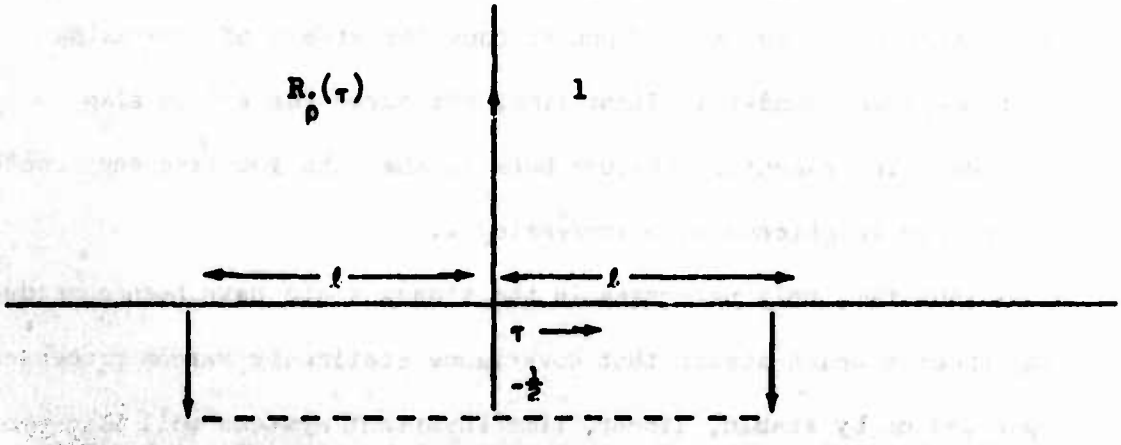


Figure 2.3. Normalized Autocorrelation Function of Doppler Data Noise Arising from White Oscillator Noise

The power spectrum of the output of the entire system, i.e., the spectrum of the range-change or integrated doppler errors, is computed as follows

$$H_2(j\omega) = \frac{1}{j\omega}$$

$$H_1(j\omega)\, H_2(j\omega) = (1 - e^{-j\omega\ell})/j\omega$$

$$|H_1(j\omega)\, H_2(j\omega)|^2 = 2\,(1-\cos\omega\ell)/\omega^2$$

Thus

$$G_{\Delta\rho}(\omega) = 2K\,(1-\cos\omega\ell)/\omega^2 \qquad (2.4)$$

Again, it is useful to inspect the behavior near $\omega=0$. Employing L'Hospital's rule

$$\frac{d(1-\cos\omega\ell)}{d(\omega^2)} = \frac{\ell\sin\omega\ell}{2\omega}$$

Re-employing on the first result,

$$\frac{d(\ell\sin\omega\ell)}{d(2\omega)} = \frac{\ell^2\cos\omega\ell}{2}$$

this last result becomes $\ell^2/2$ at $\omega=0$. Thus

$$G_{\Delta\rho}(\omega) = K\ell^2$$

$$\omega\to 0$$

Equation (4) is plotted in Figure 2.4. The plot is normalized by setting $K=1/2$ and $\ell=1$. Then to show the effect of increasing station-probe round-trip light time, the curve for $\ell=2$ is also plotted. The essential feature here is that the low frequency content is sharply heightened with increasing ℓ.

Thus far, only processes in the steady state have been considered; the theorem which states that covariance stationary random processes operated on by stable, linear, time invariant systems will also produce (in the steady-state) covariance stationary processes has

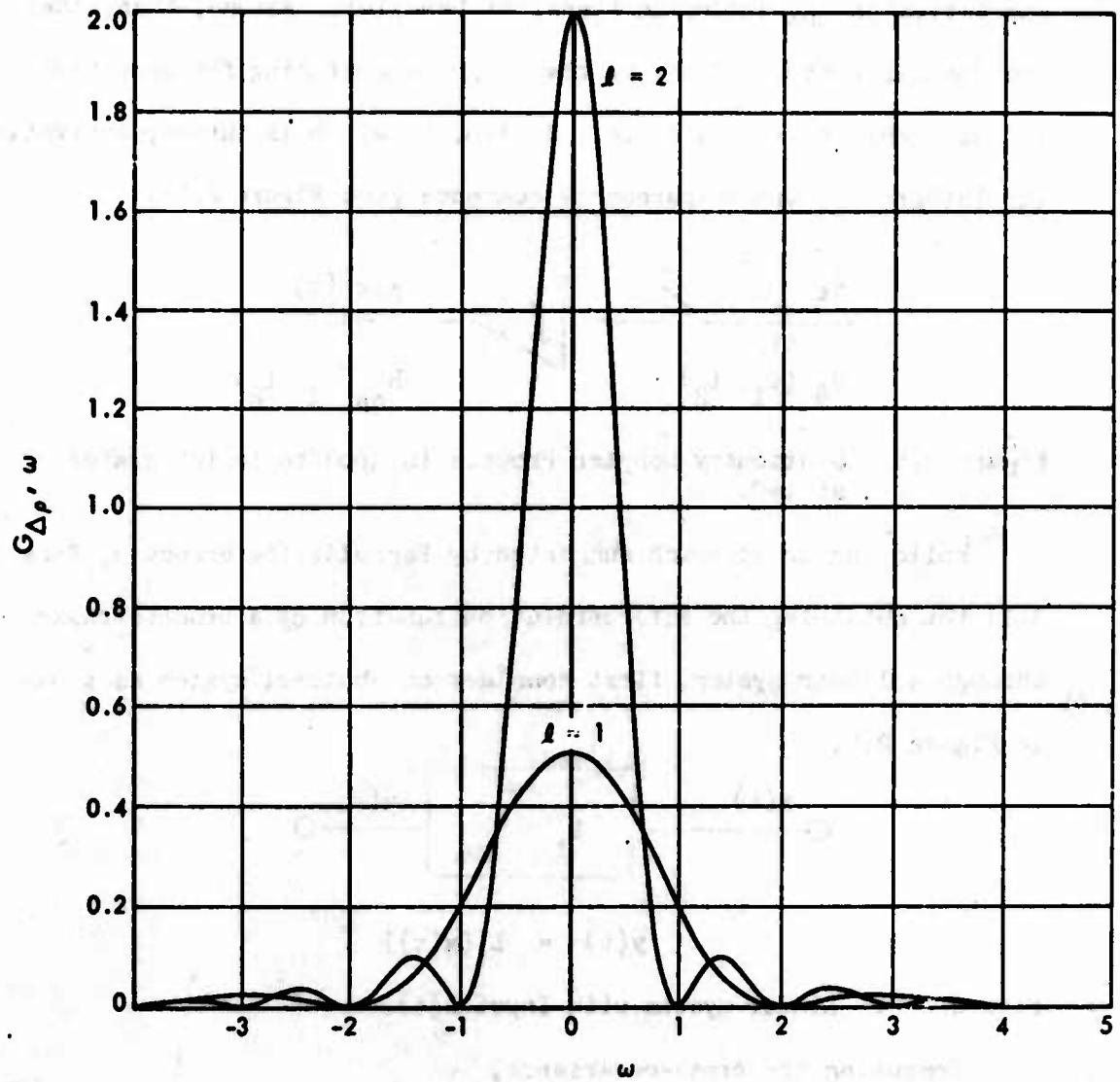


Figure 2.4. Integrated Doppler Error Power Spectral Density Arising from White Oscillator Frequency Noise.

implicitly been used. Recognizing that tracking systems operate in disjoint time segments of finite length, it is necessary to derive statistics on $\Delta\rho\epsilon$ including transient behavior. Assume, then, that the doppler portion of the system has been operating for some time and has achieved steady-state. At t=o, a switch is thrown, activating the integrator, and measurements commence (see Figure 2.5).

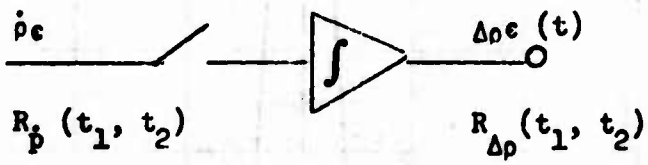
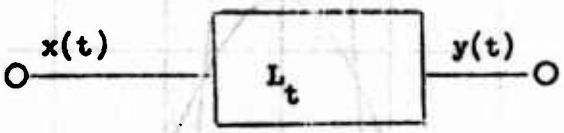


Figure 2.5. Stationary Doppler Process is Applied to Integrator at t=O.

Following an approach suggested by Papoulis (Reference 3, Page 311) for obtaining the autocorrelation function of a process passed through a linear system, first consider an abstract system as shown in Figure 2.6.



$$y(t) = L_t(x(t))$$

Figure 2.6. Linear System with Input x(t).

Computing the cross-covariance,

$$x(t_1)\, y(t) = L_t(x(t_1)\, x(t))$$

$$E\left[x(t_1)\, y(t)\right] \triangleq R_{xy}(t_1, t_2) = E\left[L_t(x(t_1)x(t)\right]$$

interchanging the order of the expectation and the system operators, the following intermediate result is obtained:

$$R_{xy}(t_1 t) = L_t(R_{xx}(t_1, t)) \qquad (2.5)$$

Leaving the abstract notation and returning to the physical problem, note that the L_t is the integral operator and R_{xx} is the instantaneous doppler error autocorrelation function given by (2.3). Thus,

$$R_{\rho\Delta\rho}(t_1, t_2) = \frac{K}{\pi}\int_0^{t_2}\left[\delta(t_1-t) - \tfrac{1}{2}\delta(t_1-t-\ell) - \tfrac{1}{2}\delta(t_1-t+\ell)\right]dt$$

$$R_{\rho\Delta\rho}(t_1, t_2) = \frac{K}{\pi}\left\{U(t_2-t_1) - \tfrac{1}{2}U(t_1-\ell)\,U(t_2-(t_1-\ell)) - \tfrac{1}{2}U(t_2-(t_1+\ell))\right\} \qquad (2.6)$$

where $U(t) = \begin{cases} 0 & t<0 \\ 1 & t\geq 0 \end{cases}$

the unit step function.

Equation (2.6) is an intermediate result; the autocorrelation of the integrated output, $R_{\Delta\rho}(t_1, t_2)$, is desired.

Reverting to the abstract system notation and continuing from Equation (2.5):

$$R_{yy}(t_1, t_2) \triangleq E\left[y(t)\,y(t_2)\right] = E\left[L_t x(t)\,y(t_2)\right]$$

Again interchanging the operators

$$R_{yy}(t, t_2) = L_t R_{xy}(t, t_2) \qquad (2.7)$$

Or, in the notation of the physical system and substituting the intermediate result (2.6) into (2.7),

$$R_{\Delta\rho}(t_1, t_2) = \frac{K}{\pi}\int_0^{t_1} U(t_2-t) - \tfrac{1}{2}U(t-\ell)\,U(t_2-(t-\ell)) - \tfrac{1}{2}U(t_2-(t+\ell))\,dt \qquad (2.8)$$

The evaluation of (2.8) is rather tedious, but straightforward. Omiting the algebraic detail, the final result becomes:

$$R_{\Delta\rho}(t_1, t_2) = \frac{K}{\pi}\left\{t_2 + (t_1-t_2)\,U(t_2-t_1)\right.$$

$$-\tfrac{1}{2}\Big[(t_1-\ell)U(t_1-\ell)U(t_2+\ell-t_1) + t_2U(t_1-(t_2+\ell)) \qquad (2.9)$$

$$+ (t_2-\ell)U(t_2-\ell)U(t_1-(t_2-\ell)) + t_1U(t_2-\ell-t_1)\Big]\Big\} \text{ for } t_1,\ t_2 \geq 0$$

This is the desired result: the autocorrelation of the integrated doppler error assuming a stationary input to the final integrator and measuring from t=0. Equation (2.9) is difficult to regard directly and arrive at any sensible interpretation. If, however, we set

$$t_1 = t_2 = t$$

a simple result for the variance of the process is obtained

$$R_{\Delta\rho\Delta\rho}(t,\ t) = \frac{K}{\pi}\Big\{t-(t-\ell)U(t-\ell)\Big\}$$

or

$$R_{\Delta\rho\Delta\rho}(t,\ t) = \frac{K}{\pi}\,t \qquad t < \ell \qquad (2.10)$$

$$\frac{K}{\pi}\,\ell \qquad t \geq \ell$$

Equation (2.10) is plotted in Figure 2.7.

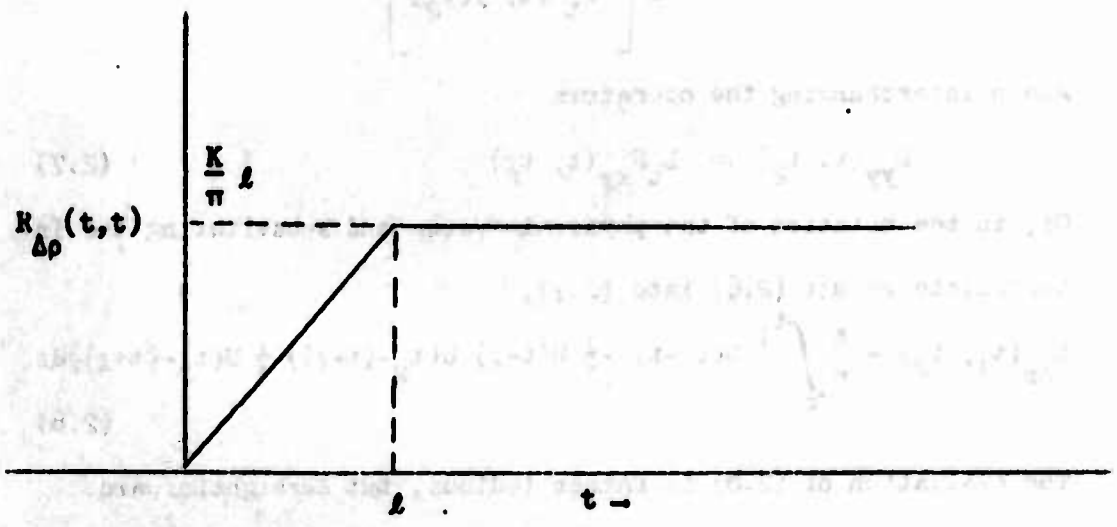


Figure 2.7. Variance of $\Delta\rho_c$ versus Tracking Time, t

the autocorrelation rises until $t_1 = \ell$ and then remains constant at .5 until t_1 nears t_2. It then rises sharply to obtain a maximum value of 1 (which checks with (2.10))and then falls to a constant value of .5 for all $t_1 \geq t_2 + \ell$. We are now in a position to make and demonstrate two assertions:

1. For $t_1 \geq \ell$, $\Delta\rho_\epsilon(t)$ is a wide sense-(or covariance-) stationary process.

In order for the assertion to be true, the process $\Delta\rho_\epsilon(t)$ for $\ell \leq t < \infty$ must have the following properties:

i. $\left|E\left[\Delta\rho_\epsilon(t)\right]\right| < \infty$

ii. $\left|E\left[\Delta\rho_\epsilon(t_1)\, p_\epsilon(t_2)\right]\right| < \infty$

iii. $E\left[\Delta\rho_\epsilon(t)\right] = m$, a constant

iv. $E\left[(\Delta\rho_\epsilon(t_1) - m)(\Delta\rho_\epsilon(t_2) - m)\right]$

$= \gamma(t_1, t_2) = \gamma(|t_1 - t_2|)$

Properties (i) - (iii) clearly obtain from what has already been said and from the fact that since the input process has mean zero and the initial conditions are zero, the output must have mean zero also. Property (iv) can be checked by re-writing the general autocovariance expression (2.9) using the constraint $t_1, t_2 > \ell$. The following result obtains:

$$R_{\Delta\rho}(\tau) = \frac{K}{\pi}\,\frac{\ell}{2}\left\{+\,\frac{\ell - |\tau|}{2}\;U(\ell - |\tau|)\right\} \qquad (2.11)$$

where $t_1, t_2 \geq \ell$

and $\tau \triangleq t_1 - t_2$

Thus $\Delta\rho_\epsilon(t)$ is a wide-sense stationary process.

Thus the variance goes linearly with time until t equals the round-trip light time (or delay, ℓ), at which point the steady-state is achieved. This system has the rather unusual feature of achieving steady-state in finite time. With this insight, return to the general expression (2.9), hold t_2 fixed and inspect the behavior versus t_1. A particular example is shown in Figure 2.8. Here $K = \pi$, $\ell = 1$ for normalization. As shown

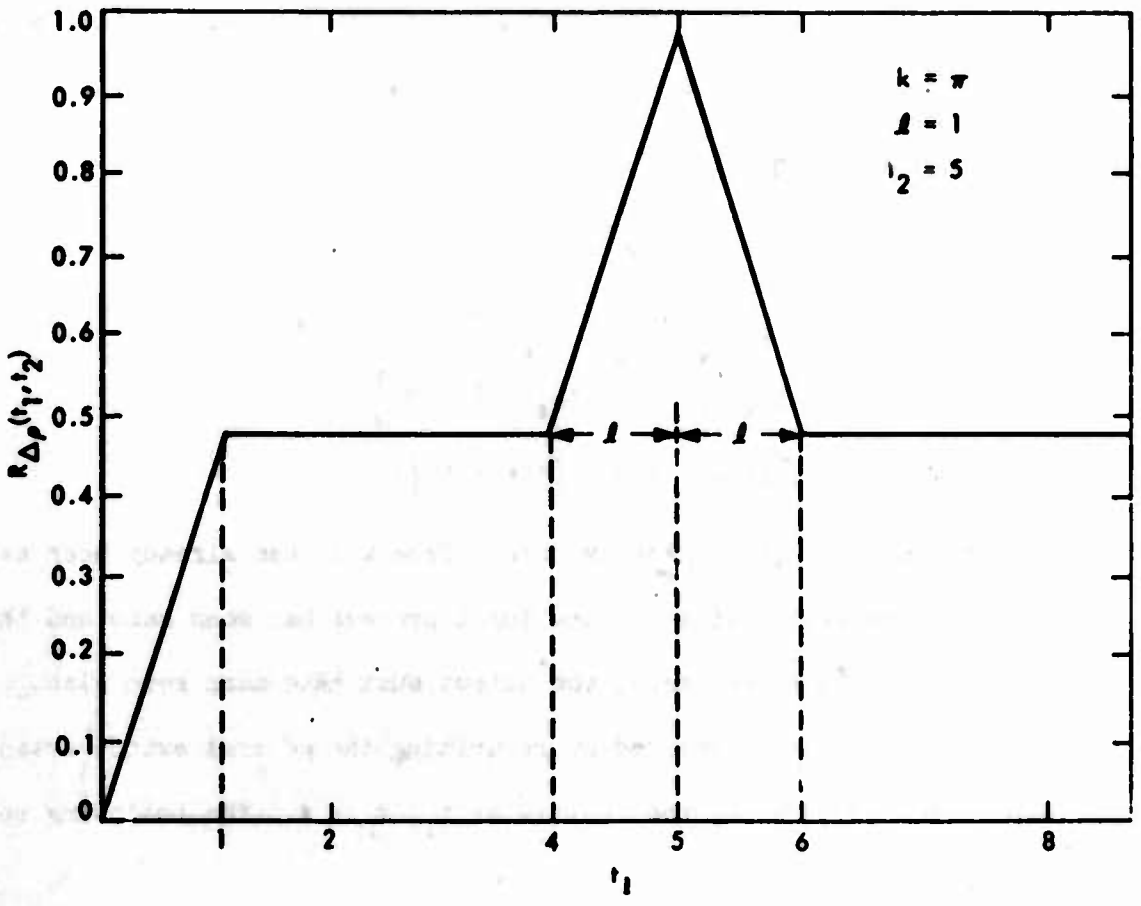


Figure 2.8. $R_{\Delta\rho}(t_1, t_2)$ Versus t_1

2. For t_1, $t_2 \geq \ell$, $\Delta\rho_e$ can be assumed to be generated by two independent processes; i.e. $\Delta\rho_e(t)$ can be written as:

$$\Delta\rho_e(t) = a + b(t) \tag{2.12}$$

where a is a mean-zero random variable of variance

$$\sigma_a^2 = \frac{k\ell}{\pi^2} \tag{2.13}$$

but acts as a bias on any one experiment, and b(t) is a mean zero wide sense stationary stochastic process, independent of a and having autocorrelation function.

$$R_b(\tau) = \frac{K}{\pi} \frac{(\ell - |\tau|)}{2} U(\ell - |\tau|) \tag{2.14}$$

The validity of this assertion is shown by using (2.13) and (2.14) in (2.12) and demonstrating the original correlation function (2.11) is obtained. Equation (2.14) is shown plotted in Figure (2.9)

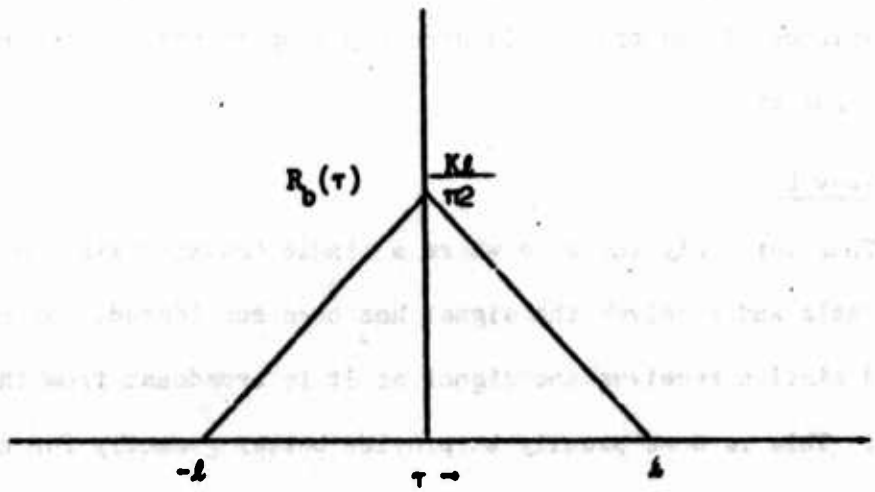


Figure 2.9. Autocorrelation of One Component of $R_{\Delta\rho}(\tau)$

Further, a stochastic process having identical first and second moment characteristics to b(t) can be generated simply by passing white noise through a smoother; i.e.

$$b(t) = \sqrt{\frac{K}{\pi}} \int_{t-\frac{\ell}{2}}^{t+\frac{\ell}{2}} w(\alpha)d\alpha \tag{2.15}$$

where $w(\alpha)$ is a white noise process of unity power spectral density. In summary,

1. the output of the integrated doppler system due to white master oscillator frequency variations is described in general by a non-stationary process, but this process reaches a steady-state in finite time.

2. If the output is sampled only after steady-state has been achieved (ℓ seconds after initiation), the error can be described for visualization or perhaps simulation purposes as having been generated by a bias (random between experiments and with mean zero, but constant during a single experiment) and a white noise smoother as described by (2.15). Each of these sources contribute equally to the variance. The variance of the process is directly proportional to the round-trip light time.

Three-way Data

Thus far, only the case where a single tracking station both broadcasts and receives the signal has been considered. Often, a second station receives the signal as it is broadcast from the spacecraft. This is done usually to provide better geometry for near-Earth tracking (less then one lunar radius) since the stations can be separated by as much as an Earth diameter. This configuration is the so-called three-way data taking mode, because there are three distinct participants. This new system can be modeled in a manner analagous to Figure 2.1. This is shown as Figure 2.10. Here, independent oscillators are needed.

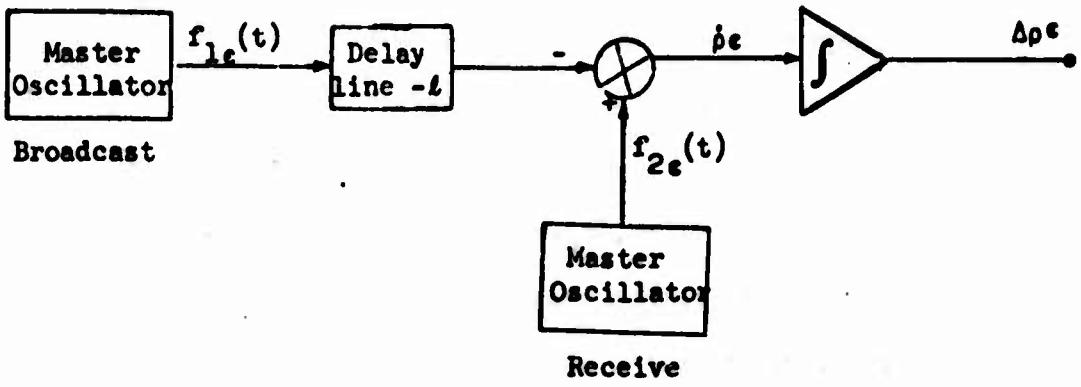


Figure 2.10. Block Diagram of Three-Way Doppler System

and there can be no feed forward of the input errors to ultimately cancel and keep the output a stationary process; the system is now unstable. In addition, the original assumption that $f_\epsilon(t)$ contained no bias or offset errors cannot be so blithely made; for three-way data, an oscillator bias directly produces a bias in the doppler, $\dot{\rho}$.

Even so, suppose the estimation process to be not sensitive to the bias and only the stochastic errors in the data to be of real concern. It is desired then, to calculate the effects of a white noise oscillator input on $\Delta\rho$ as was done before. The calculations proceed similarly but are much simpler; therefore only the results are stated. If both oscillators have constant spectral density, K, the autocorrelation function on the doppler is simply

$$R_{\dot{\rho}}(\tau) = \frac{K}{\pi}\,\delta(\tau) \qquad (2.16)$$

Thus, if the output were to be smoothed over short time intervals ($<\ell$) so that we may meaningfully speak of variance, the variance on the three-way data is identical to that of its two-way counterpart (compare with(2.3)).

Often, analyses of doppler errors proceed no further than this

since for conceptual purposes the data can be considered as direct $\dot{\rho}$ measurements by suitably sampling and differencing the output of the integrator. The importance of the negative correlation spikes in (2.3) can only be underlined, however, by passing the three-way data through the final integration stage and comparing the results there.

Clearly (2.16) indicates that the $\dot{\rho}_\epsilon$ is a simple white noise process, which when integrated yields a Wiener process, i.e:

$$R_{\Delta\rho}(t_1, t_2) = \frac{K}{\pi} t_1 \qquad t_1 < t_2$$

$$= \frac{K}{\pi} t_2 \qquad t_2 \leq t_1$$

or the variance

$$R_{\Delta\rho}(t,t) = \frac{K}{\pi} t \qquad \text{for all } t \geq 0$$

Thus while the two-way error is limited by the round-trip light time, ℓ, the three-way error variance continues to build throughout the tracking interval. For a typical light time of a 1000 sec. and a tracking interval of twelve hours, the relative variance of the oscillator noise on three-way as to two-way data would be in a ratio of approximately 30 to 1.

Numerical Results

Doppler tracking data was obtained from the Pioneer VI spacecraft at one second sampling rates. Successive samples were differenced and divided by the count time (in this case unity) to produce smoothed $\dot{\rho}$ samples. The autocorrelation function of these data was taken with results as shown in Figure 2.11.

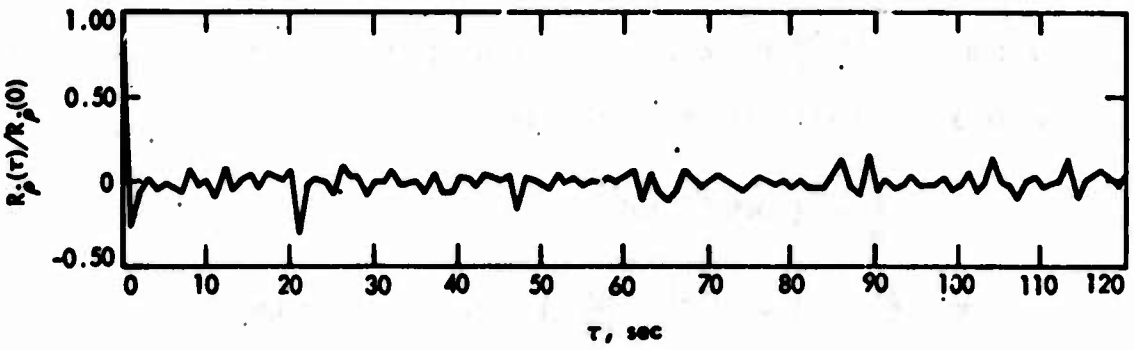


Figure 2.11. Normalized Autocorrelation Function for Pioneer VI Doppler Tracking Data.

A negative spike is expected in the function at a lag near the round-trip light time, which in this case was 21.3 sec. as predicted by the theory (c.f. Equation (2.3)). This expectation is clearly rewarded as seen in Figure 2.11. The negative spike at one lag time is due to another phenomenon, receiver phase jitter, and can also be explained by a simple model (see Curkendall, Reference 4), but is of no consequence to the present pursuit. It does serve to separate the receiver jitter contribution to the high frequency noise from that arising from the oscillator. This separation yields that

$$\frac{K}{\pi} = (.034)^2 \quad \text{hz}^{2*} \qquad \left(\frac{K}{\pi} \text{ is defined by (2.3)}\right)$$

*Here a cycle is one cycle of integrated doppler which at the S band broadcast frequency of 2200×10^6 hz is equivalent to 6.5 cm. in range change.

This would arise from an rms deviation in the master oscillator, $\frac{\Delta f}{f}$ of $1.4 \times 10^{-11} \frac{hz}{hz}$ for a 1 sec. averaging time. Other similar experiments yielded deviations as high as

$$\frac{K}{\pi} = (.045)^2 \quad hz^2$$

or $\frac{\Delta f}{f} = 2 \times 10^{-11} \quad \frac{hz}{hz}$ 1 sec. smoothing time.

Employing this later figure in the integrated doppler variance (Equation 2.10),

$$R_{\Delta\rho} = (.045)^2 \ell$$

assuming ℓ to be 20 min. and converting to metric units,

$$\sigma_{\Delta\rho} \triangleq \sqrt{R_{\Delta\rho}(t,t)} = (.045) \frac{cy}{\sqrt{sec}} \times \sqrt{1200 \text{ sec}} \times 6.5 \frac{cm}{cy}$$

$$= 10 \text{ cm.}$$

The total integrated doppler error from all sources (see the discussion in Appendix II) is approximately 150 cm; thus the white noise component from rubidium oscillators contributes but a small part of this total error. But the longer term drifts also characteristic of these same oscillators can be shown to be a major contributor. A reasonable drift over a 12 hr pass is 5 parts in 10^{12}, much smaller than the measured short-term fluctuations of 2 parts in 10^{11}. Its effect can be computed in the following manner:

Observe that $\Delta\rho_e$ can be written directly in terms of the $f_e(t)$ of Figure 2.1 and an integral operator

$$\Delta\rho_e = k \int_0^T \left[f_e(t) - f_e(t-\ell) \right] dt \qquad (2.17)$$

where the k is inserted to change to metric units and T is the total tracking interval (e.g. 12 hr.). Noticing that cancellation takes

place, we can re-write (2.17) as

$$\Delta\rho_\epsilon = k\left[\int_{T-\ell}^{T} f_\epsilon(t)dt - \int_{-\ell}^{0} f_\epsilon(t)dt\right] \qquad (2.18)$$

If f_ϵ is a slowly varying quantity (i.e. it changes little over a period of ℓ seconds)

$$\Delta\rho_\epsilon \approx k\ell\left[f_\epsilon(T) - f_\epsilon(0)\right] \qquad (2.19)$$

Using the postulated drift of 5 parts in 10^{12} or .011 hz at S band (2200 Mhz)

$$\Delta\rho_\epsilon = 80 \text{ cm}$$

Thus it is shown, that while the high frequency variations from the oscillator are much larger in amplitude than the assumed longer term drift, because of the cancellation afforded by the feed forward loop of Figure 2.1, their errors are much less important than those arising from the slower (and invisible!) drifts that take place over a longer time scale. Moreover, the effects of the high frequency components grow with $\sqrt{\ell}$, while the low frequency variations go with ℓ. For outer planet missions, where the light time can be on the order of 8 hr., the current rubidium oscillators cannot perform adequately (they are marginal at Mars distances).

CHAPTER III.

A GEOMETRICAL INTERPRETATION OF EARTH-BASED TRACKING DATA FOR AN INTERPLANETARY SPACECRAFT

In this chapter our attention is turned toward using the measurements of doppler and range to determine the orbit of a spacecraft. These two measurement types, implemented as discussed in the Appendix, are universally applicable to all phases of space-flight past the initial powered flight-Earth orbit: trans-lunar, lunar orbit and lunar lander, interplanetary, and soon, planetary operations both orbiter and lander. Consideration here, however, will be restricted to the interplanetary portion of flight.

This phase of the mission is characterized by long tracking arcs, rather weak gravitational fields and hence seemingly monotonous tracking geometries. The lack of rapidly changing geometry such as is present, for example, in a lunar orbiter problem, places the most stringent accuracy demands upon the tracking systems and physical models so that the inherent information concerning the orbit that is present can be observed with the requisite fidelity. Conversely, this same simple geometry can be described by models transparent enough to allow the critical (relevant) elements of the problem to be discerned. The analysis to be presented here has resulted from the efforts of several investigators of which this author is only one. In 1965, J. Light (Reference 5) first observed that there was an important distinction between trans-lunar and interplanetary flights due to the

distances involved. In the former, the angular diameter of the Earth as seen from the moon ($\sim 2^{\circ}$) was sufficient so that multiple tracking stations, viewing the trajectory from slightly different aspects could provide enough position parallax to fully determine the orbit over very short time intervals. At planetary distances, this position parallax is of no consequence, but Light noticed that the rotation of the tracking station about the Earth's spin axis impressed a diurnal signature on the doppler tracking data whose amplitude and phase were of manifest importance. Hamilton and Melbourne (Ref. 2) formalized these observations and put the analysis in a form that showed the dominant factor limiting orbit determination accuracy was the accuracy in the knowledge of the station locations themselves. Later, this author extended the Hamilton/Melbourne analysis, which considered only three spacecraft state parameters, to include the remaining variables and produced a full-rank (six parameter) description of orbit determination accuracy (Ref. 6). Perhaps the most extensive analysis to date was published in 1969 by this author and S. R. McReynolds (Ref. 7). In this chapter we shall recall that analysis, its major results, and provide the background for additional results in subsequent chapters.

The Doppler Data Equation

Consider Figure 3.1 which shows the Earth-relative geometry present when tracking a spacecraft from a tracking station displaced from the Earth's spin axis by the distance r_s (it is also elevated above the Earth's equatorial plane, but this is of no consequence). If $r \gg r_s$ and the spacecraft were not moving with respect to the Earth, the range-rate (or doppler) would entirely be composed of a harmonic

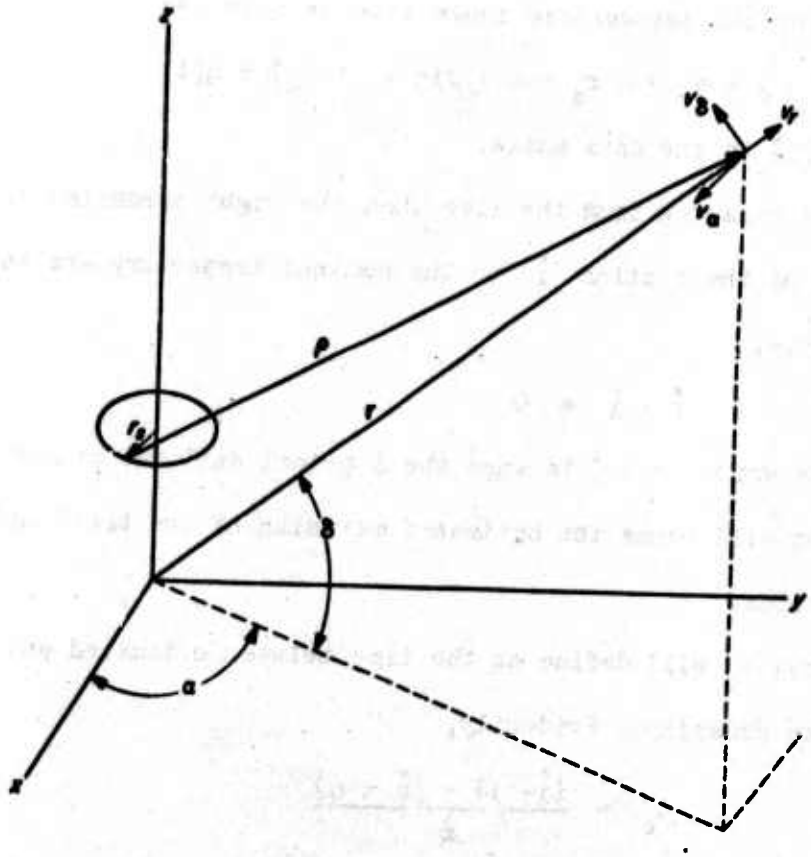


Figure 3.1. Geocentric Equatorial Coordinate System.

component of amplitude

$$\omega_m r_s \cos \delta$$

where ω_m is the Earth's rotation rate and all other variables are defined in Figure 3.1. If we add the radial velocity, the first approximation for the topocentric range-rate is obtained:

$$\dot{\rho} = v_r + \omega_m r_s \cos \delta \sin \omega_m (t-t_o) + n(t) \qquad (3.1)$$

where n(t) is the data noise.

Let t be measured from the time when the right ascension of the probe ($\hat{\alpha}$) and of the station ($\hat{\lambda}$) on the nominal trajectory are in coincidence, i.e.,

$$\hat{\alpha} - \hat{\lambda} = 0$$

In other words, $t = 0$ is when the a priori estimate of the probe's position will cross the estimated meridian of the tracking station's coordinates.

Thus t_o will define as the time between estimated and actual meridian crossing. Evidently,

$$t_o = \frac{(\hat{\lambda} - \lambda) - (\hat{\alpha} - \alpha)}{\omega}$$

where

$$= \frac{\Delta\alpha - \Delta\lambda}{\omega}$$

$$\Delta\alpha \triangleq \alpha - \hat{\alpha}$$

$$\Delta\lambda \triangleq \lambda - \hat{\lambda}$$

Equation 3.1 becomes

$$\begin{aligned} \dot{\rho} &= v_r + \omega r_s \cos \delta \sin (\omega_m t - \Delta\alpha + \Delta\lambda) + n(t) \\ &= v_r + \omega r_s \cos \delta \left\{ \sin \omega t \cos (\Delta\alpha - \Delta\lambda) - \cos \omega_m t \sin (\Delta\alpha - \Delta\lambda) \right\} \\ &\quad + n(t) \end{aligned} \qquad (3.2)$$

If the a priori estimate of the station's longitude and the probe's

right ascension are reasonably accurate, then

$$|\Delta\alpha - \Delta\lambda| \ll 1$$

and (3.2) can well be approximated by

$$\dot{\rho} = v_r + \omega r_s \cos\delta \sin\omega t + \omega r_s \cos\delta\,(\Delta\lambda - \Delta\alpha)\cos\omega t + n(t) \tag{3.3}$$

Or, re-writing in terms of the unknown coefficients,

$$\dot{\rho} = a + b \sin\omega t + c \cos\omega t + n(t)$$

Thus, over a time interval when v_r, δ, and α can be considered constant, the rank of the doppler information is at most three. Equation (3.3) was derived in Reference 2.

Suppose we now consider the three variables in a first-order time series expansion, i.e.,

$$v_r(t) = v_r + a_r t$$
$$\delta(t) = \delta + \dot{\delta} t$$
$$\alpha(t) = \alpha + \dot{\alpha} t$$

Substituting these expressions into (3.3) and using the following approximation

$$\cos(\delta + \dot{\delta}t) \approx \cos\delta - \sin\delta\,\dot{\delta}t$$

we obtain:

$$\dot{\rho} = a + b \sin\omega t + c \cos\omega t + dt + et \sin\omega t + ft \cos\omega t + n(t) \tag{3.4}$$

where

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} \triangleq \begin{bmatrix} v_r \\ \omega r_s \cos\delta \\ \omega r_s \cos\delta\,(\Delta\lambda - \Delta\alpha) \\ a_r \\ -\omega r_s \sin\delta\,\dot{\delta} \\ -\omega r_s \left\{(\Delta\lambda - \Delta\alpha)\sin\delta\,\dot{\delta} + \dot{\alpha}\cos\delta\right\} \end{bmatrix}$$

With the exception of the d term, all of the coefficients are written explicitly in terms of the 6 unknown state parameters; r, δ, α, v_r, $\dot{\delta}$, $\dot{\alpha}$. This d, or acceleration, term is composed of two major effects:

1. The component of the Sun's gravitational acceleration along the line-of-sight to the probe.
2. The so-called "centrifugal acceleration" arising from a non-totally outward directed spacecraft velocity.

We shall leave the first of these implicit for the time being but re-write d as

$$d = a_g(r, \delta, \alpha) + r(\dot{\alpha}^2 \cos^2 \delta + \dot{\delta}^2)$$

where a_g is the gravitation component of the acceleration and the second term is the familiar $\omega^2 r$ from elementary dynamics.

Figure (3.2) displays the doppler signatures assuming the three parameter model (a) and then the six-parameter model (b). In (b) the signature from the second day's tracking is shifted to the left by 24 hrs., so that the effects of the presence of d, e, and f can be emphasized.

The Range Data Equation

At $t = 0$, the spin axis of the Earth, the station vector $\bar{r}_{st}$, and the probe vector, $\bar{r}_o$, are co-planar as shown in Figure 3.3. Invoking the approximation

$$\frac{\hat{\rho}}{|\hat{\rho}|} \approx \frac{\hat{r}}{|\hat{r}|}$$

Then

$$\rho(t_o) = r(t_o) - |\bar{r}_{st}| \cos(\delta - \theta) + n(t_o) \qquad (3.5)$$

It can easily be shown that range data taken at other times, t_i,

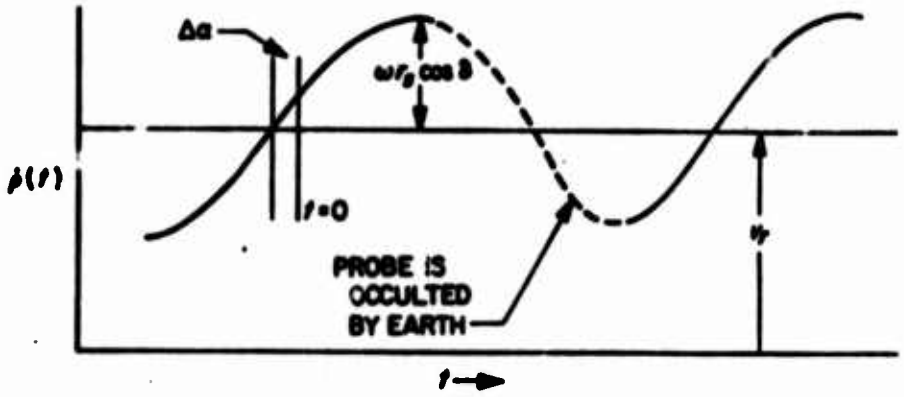
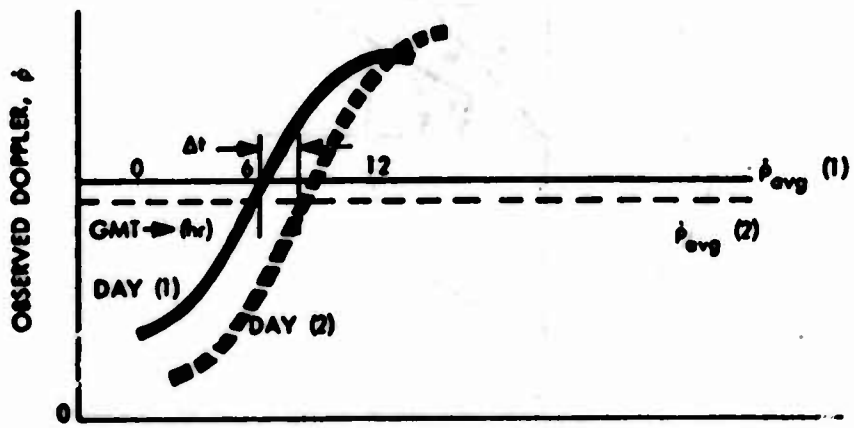


Figure 3.2 (a). Doppler Signature Using First Three Terms.



CHANGE IN { AVERAGE ρ̇ ; TIME OF MERIDIAN CROSSING ; AMPLITUDE OF HARMONIC COMPONENT } IS PROPORTIONAL TO { ACCELERATION ; RIGHT ASCENSION RATE ; DECLINATION RATE }

DOPPLER SIGNATURES ON TWO SUCCESSIVE DAYS. FOR THE SITUATION SHOWN, THE SPACECRAFT ACCELERATES OUTWARD, MOVES COUNTER-CLOCKWISE WITH RESPECT TO THE EARTH, AND EXPERIENCES A DECREASE IN THE DECLINATION MAGNITUDE.

Figure 3.2 (b). Doppler Signature Using Six Parameter Expansion.

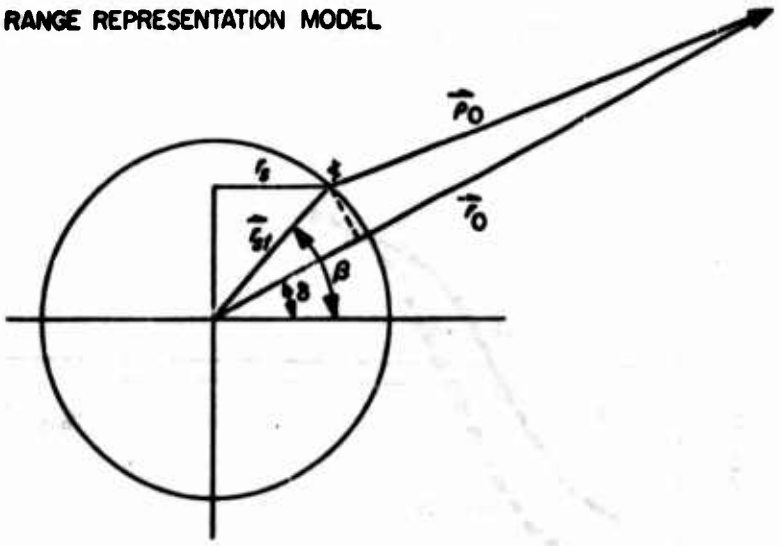


Figure 3.3. Center of Pass Range Geometry.

during the pass can be referred to t_o by adding to $\rho(t_i)$ the integral of the measured doppler data between points t_i, t_o nearly without additional error. This is because, over short periods of time a modern doppler system can measure range change much more accurately than can the ranging system. A pass of range data, then, (if it accompanies a pass of doppler data) can be compressed to a single point by referring all points to t_o and performing a simple algebraic average as a first filtering operation.

The Concept on Data Compression as an Accuracy Analysis Tool

The normal procedure in an estimation problem is to form the partials of the data with respect to the state of the system and estimate that state from the raw data in a single operation. However, note from (3.4), that the data is written as a linear function of the parameters a - f. Thus, if the system state is first thought of as being these parameters, the accuracy of their estimation will not be dependent upon the nominal values since those nominal values do not appear in the data partials (linearity). This is useful. A pass or more of doppler data (perhaps including as many as 1000 observations) can be considered as equivalent to the observation of the six parameters, a - f. Moreover, since the accuracy of observation of these six parameters is not a function of probe geometry, the computation of how well they can be observed (for a given data noise model) need be performed only once.

Calculating the accuracy of the estimates of these parameters for a given length of tracking data (perhaps a twelve hour pass) can easily

be accomplished with any of the estimation algorithms detailed in the Appendix. Assume this has been done and the resultant covariance of errors is found to be Λ_a. Estimation of the state parameters r, δ, α, ..., $\dot{\alpha}$ is accomplished by taking the partials of a - f with respect to the state parameters and using any standard estimation algorithm. Using, the normal equation formulation, the covariance of estimation errors is given by

$$\Lambda_r = (A^T J_a A)^{-1} \tag{3.6}$$

where

$$J_a \triangleq \Lambda_a^{-1}$$

and A, the partial matrix, is readily seen to be,

$$A = \begin{array}{c} \\ a \\ b \\ \\ c \\ d \\ e \\ f \end{array} \begin{array}{c} \begin{array}{ccccccc} r & & \delta & & \alpha & v_r & \dot{\delta} \quad \dot{\alpha} \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -\omega r_s \sin\delta & 0 & 0 & 0 & 0 \\ 0 & 0 & -\omega r_s \cos\delta & 0 & 0 & 0 \\ \frac{\partial ag}{\partial r}+\dot{\alpha}^2\cos^2\delta+\dot{\delta}^2, & \frac{\partial ag}{\partial \delta}-2r\dot{\alpha}^2\cos\delta\sin\delta, & \frac{\partial ag}{\partial \alpha} & 0 & 2r\dot{\delta} & 2r\cos^2\delta(\dot{\alpha}) \\ 0 & -\omega r_s \dot{\delta}\cos\delta & 0 & 0 & -\omega r_s \sin\delta & 0 \\ 0 & \omega r_s \dot{\alpha}\sin\delta & \omega r_s \dot{\delta}\sin\delta & 0 & 0 & -\omega r_s \cos\delta \end{bmatrix} \end{array} \tag{3.7}$$

The gravitational partials have recently been put in compact analytical form by Ondrasik (Ref. 8). They are:

$$\frac{\partial ag}{\partial r} = \frac{\mu}{r_p^3}(2 - 3\sin^2\varepsilon)$$

$$\frac{\partial ag}{\partial \delta} = \mu r_e \cos\sigma\left[\left(\frac{1}{r_p^3} - \frac{1}{r_e^3}\right) - 3\frac{r}{r_p^5}(r - r_e\cos\chi)\right]$$

$$\frac{\partial a_g}{\partial \alpha} = \mu r_e \cos\delta \cos\delta_s \sin(\alpha - \alpha_s)\left[-\left(\frac{1}{r_p^3} - \frac{1}{r_e^3}\right) - 3\frac{r}{r_p^5}\right.$$

$$\left.(r - r_e \cos\chi)\right] \qquad (3.8)$$

where

μ = GM of the Sun

r_p = Sun - probe distance

ξ = Earth-probe-Sun angle

r_e = Sun-Earth distance

$\cos\sigma = \cos\delta \cos\delta_s \cos(\alpha - \alpha_s) + \cos\delta \sin\delta_s$

χ = Sun-Earth-probe angle

δ_s, α_s = geocentric declination, right ascension of the Sun

If it is desired to study the effects of augmenting the doppler data with ranging, the averaged effective ρ at t_o is simply used directly by augmenting both the J_a and A matrices. That is

$$J_a' = \begin{bmatrix} \sigma_\rho^{-2} & 0 \\ 0 & J_a \end{bmatrix} \qquad (3.9)$$

and

$$A' = \begin{bmatrix} 1 & -|\bar{r}_s t|\sin(\delta - \rho) & 0 & 0 & 0 & 0 & 0 \\ \hline & & & A & & & \end{bmatrix} \qquad (3.10)$$

Performing the full analysis as indicated by (3.6) will yield an estimate of the covariance of the state estimation errors for the particular geometry used in the evaluation of (3.7). It is useful, however, as a first approximation, to inspect the behavior of J_a^{-1} versus tracking time and relate, on an analytic basis, what is now considered

to be the errors in the data to errors in the state estimates. Figure 3.4 charts the variance in the observation of the parameters a - f versus the number of days tracking time (assuming a nominal 12 hrs./day tracking from a single station and an equivalent uncorrelated doppler data error of 1 mm/sec. for a one minute sample). It is seen that the variance of the a - c parameters goes roughly with 1/T, but the variance in d - f, because of the presence of the secular t term in the data equation goes with $(1/T)^{-\frac{3}{2}}$.

Returning now to the data partial matrix, 3.7, it can be observed that three of the state parameters, δ, $\dot{\delta}$ and $\dot{\alpha}$, are involved in more than one of the data parameters. For short tracking spans (a few days) and notably when the nominal $\dot{\delta}$ and $\dot{\alpha}$ are small, the information concerning each of the state parameters comes largely from a single data parameter. These main partials are underlined in 3.7. To the extent that this first approximation is correct, analytic expressions for the state accuracies as a function of probe geometry can be obtained through the information given in Figure 3.4. These are readily seen to be:

$$
\begin{aligned}
\sigma_r &= \sigma_d \Big/ \left(\frac{\mu}{R^3}\left[3\cos^2 \text{EPS} - 1.\right] + \dot{\alpha}^2\cos^2\delta + \dot{\delta}^2\right) \\
\sigma_\delta &= \sigma_b/\omega\, r_s\, |\sin\delta| \\
\sigma_\alpha &= \sigma_c/\omega\, r_s \cos\delta \qquad (3.11) \\
\sigma_{v_r} &= \sigma_a \\
\sigma_{\dot{\delta}} &= \sigma_d/\omega\, r_s |\sin\delta| \\
\sigma_{\dot{\alpha}} &= \sigma_f/\omega\, r_s \cos\delta
\end{aligned}
$$

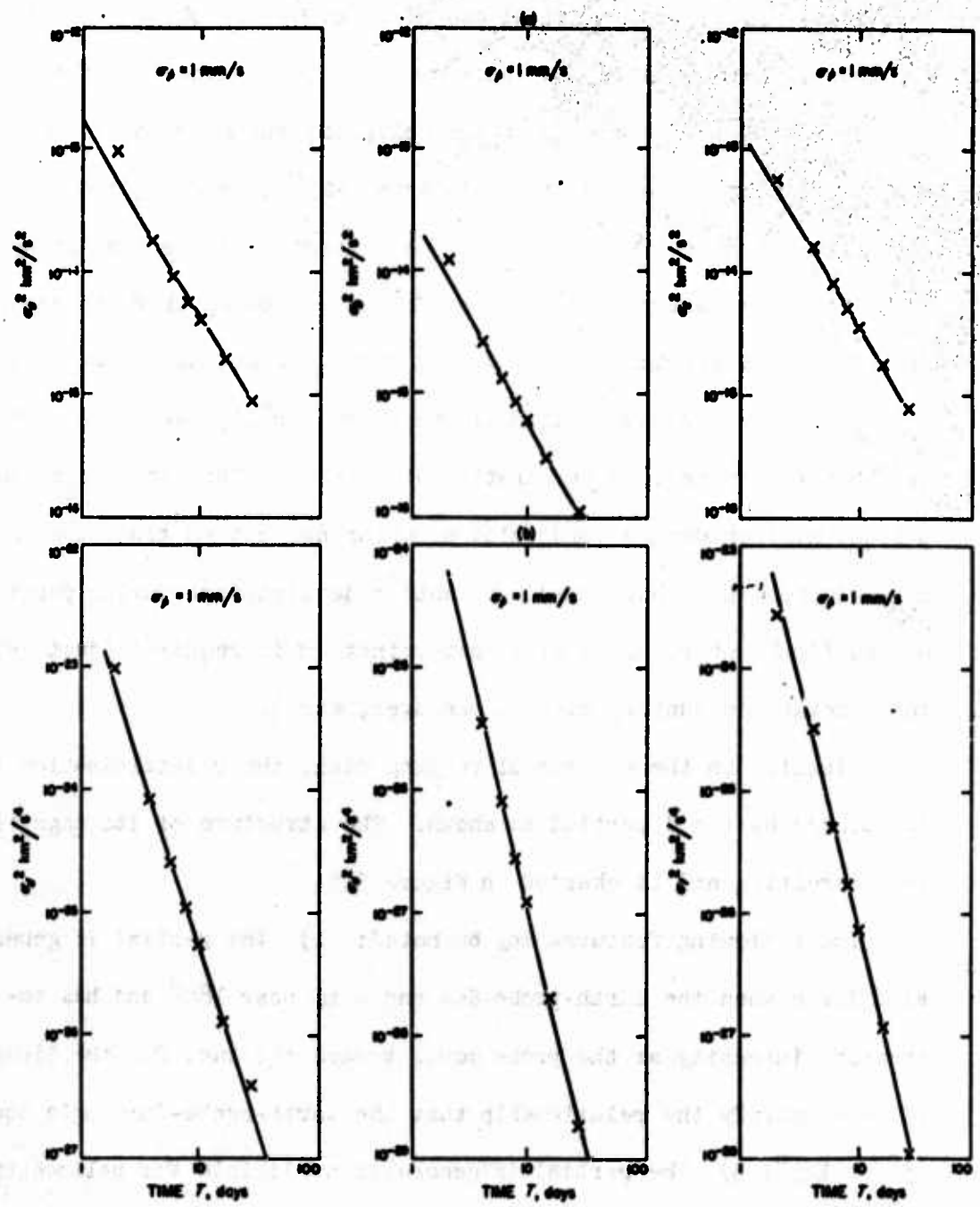


Figure 3.4. Estimation Variances $\sigma_a^2 - \sigma_f^2$ versus Tracking Time, T.

Some important geometrical dependencies can be observed. Since all spaceflights to date and those contemplated in the near future are restricted to lie nearly in the ecliptic, the geocentric declination is roughly contained in the interval -23° to $+23^\circ$ (from the obliquity of the ecliptic) and the cos δ term is always near unity. Thus, to the accuracy of the approximation now being used, the accuracies of the determinations of α, v_r, and $\dot{\alpha}$ are not dependent on the probe geometry. However, $|\sin \delta|$ can become small, even zero, and prohibit the accurate determination of δ and $\dot{\delta}$. This imposes an important mission design constraint of ensuring that no trajectories are selected which have small geocentric declinations during portions of the flight where rapid orbit determination is required (just prior to planetary encounter, after a maneuver, etc.).

Finally, in the absence of ranging data, the r determination is influenced by the d partial as shown. The structure of its magnitude is interesting and is charted in Figure 3.5.

The following features may be noted: 1) The partial is generally large when the Earth-probe-Sun angle is near 180° and has increasing intensity as the probe moves toward the sun; 2) the lines of zero satisfy the relationship that the Earth-probe-Sun angle equals 55° or 125°; 3) The partial is generally negligible for heliocentric ranges greater than 1.5 a.u. and would also be less important for a Mars transfer than for a Venus transfer.

Of more practical importance than lines where

$$\frac{\partial a_g}{\partial r} = 0$$

would be occasions when

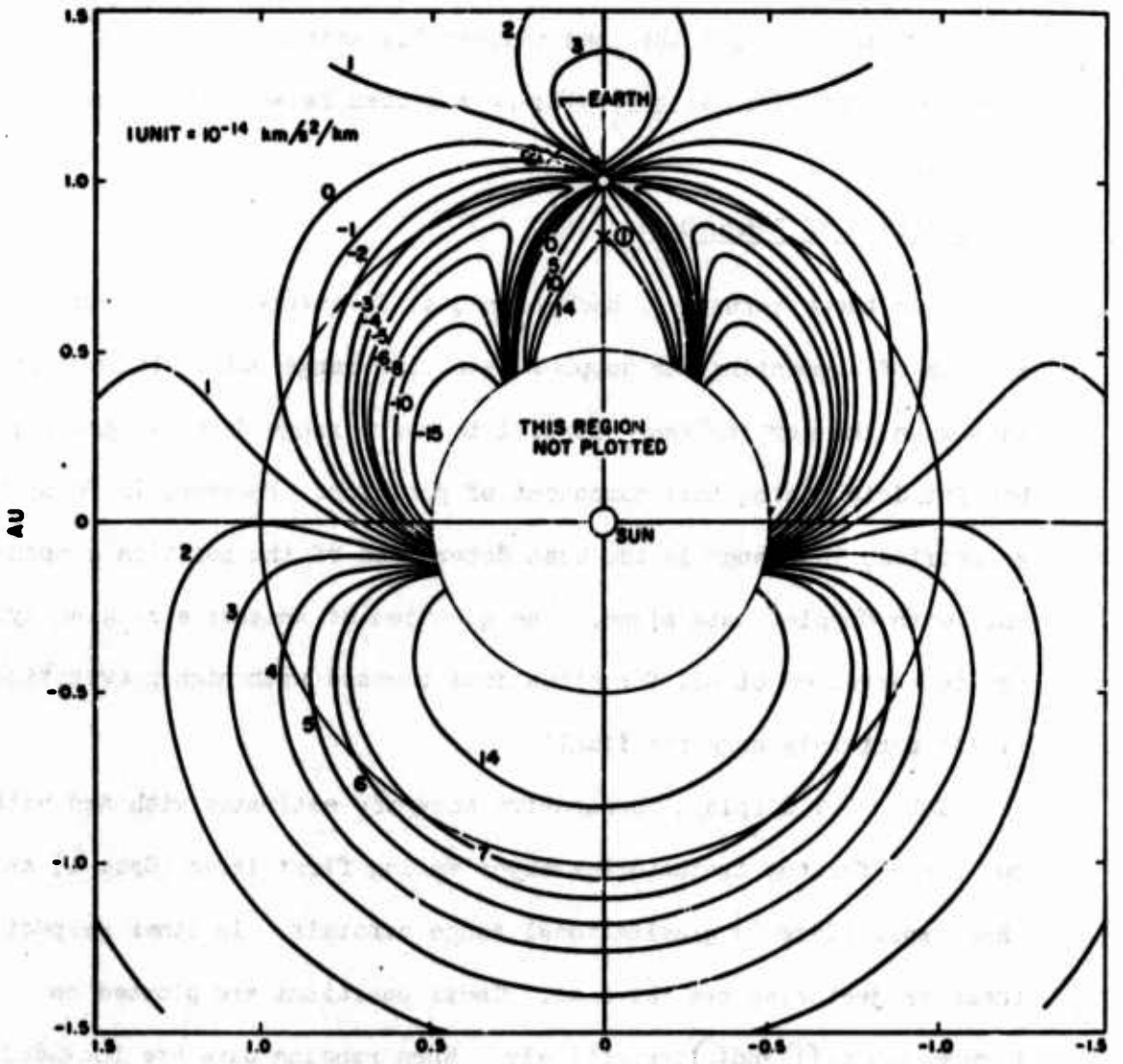


Figure 3.5. Map of Partial Derivative of the Differential Gravitational Acceleration with Respect to Geocentric Range.

$$\frac{\partial a_g}{\partial r} + \frac{\partial a_c}{\partial r} = 0$$

In this circumstance, the range could not be determined by doppler alone. Since $\frac{\partial a_c}{\partial r} \geq 0$ this can conceivably arise only when $\frac{\partial a_g}{\partial r} \leq 0$, which from the chart is generally in the area between the Earth and the Sun.

The Addition of Range Data

With these results as background, it is interesting to chart the effects of augmenting the doppler data with range data. It is clear that when the doppler range partial is weak, range data can greatly benefit determining this component of position. However, in favorable geometries, the range is the best determined of the position components with doppler data alone. The question of whether a ranging system is warranted at all for situations blessed with high gravitational fields certainly suggests itself.

Table 3.1 displays comparative accuracy estimates with and without range for the trajectories experiencing first large (Case 1) and then small (Case 2) gravitational range partials. In other respects these trajectories are the same. Their positions are plotted on Figure 3.5 as ① and ② respectively. When ranging data are included, reliance can no longer be made on the approximation that a single data term determines a single state parameter. There are now important relationships between all the partials. The data in Table 3.1 was therefore generated via consideration of the full partial matrix and algorithms similar to (3.6).

Several principles are illustrated with these data:

(1) The estimate of range, of course, improves down to the accuracy of the ranging itself in all cases.

(2) Even when the cross-velocities are zero, the estimate of v_α is improved by adding range data. The improvement is slight for the EPS = 53° case; it is quite significant (a factor of 2.2 with two days of data) for the EPS = 180° case. The explanation is that when the cross-velocities are low (or zero) the earth-probe acceleration is caused strictly by the gravitational fields and the uncertainty in the geocentric range is the principle reason for a priori uncertainty in what the acceleration will be. Conversely, supplying this range to high accuracy will determine the geocentric acceleration. Because the d (or acceleration) term is highly correlated with f (the term that determines v_α), an improvement in knowledge of d brings with it an improvement in f, and the accuracy of v_α is consequently improved. In fact, if d is perfectly known, the uncertainty in f will improve by a factor of 2.2, the same improvement already observed in v_α for the 180° case. The full 2.2 improvement is not experienced for EPS = 53° because, in the vicinity of the sun (< 1.5 AU), a low $\partial a/\partial r$ really implies not that va_g is small but that it is perpendicular to the line of sight. Thus uncertainties in the cross positions produce significant variations of a_g, and ranging simply will not remove this uncertainty.

(3) No corresponding improvement occurs in v_δ when the a prior v_δ is zero. Here the e term, which determines v_δ is uncorrelated with d and the argument just stated does not apply.

(4) When the cross-velocity is non-zero, the addition of range

data significantly improves the determination of this velocity. The reason here is clear. Supplying range data makes the entire power of the d term available to improve the velocity estimation. This improvement occurs only in a single direction - along the a priori cross-velocity itself; e.g., notice that for the case where both v_α and v_δ are 3 km/s, little improvement occurs in either individual direction - the improvement is experienced in the direction 45° from the axes we have chosen. Coupled together, Items (2) and (4) explain the often observed but previously curious result that the addition of range data considerably increases the rapidity of orbit determination following a midcourse correction where only the velocities have been corrupted.

(5) Similar computations were made to demonstrate the dependence of 1 - 4 on the precision of the range data. The velocity improvement was found to be unrelated, however, as long as the measurement was considerably (X10) more precise than the accuracy of range determined from Doppler alone.

In summary, range data improves the accuracy of a single component of position as expected and, in addition, one component of cross-velocity. The former is strictly a function of the measurement itself, the latter depends upon the magnitude and direction of the cross-velocity.

Table 3.1. The Influence of Range Data on Orbit Accuracy

CASES			2 days of data σ_r, km	σ_{n2}, m/sec	σ_{n_s}, m/sec
			EPS = 180° ($\partial a_g/\partial r$ = 14 x 10^{-14} sec^{-2})		
v_r	= 0	No ranging	9.9	0.13	0.23
v_δ	= 0	Ranging	0.01	0.058	0.23
v_α	= 3	No ranging	193.	0.11	0.23
① v_δ	= 0	Ranging	0.01	0.0027	0.23
v_α	= 3	No ranging	373.	0.11	0.19
v_δ	= 3	Ranging	0.01	0.11	0.11
			EPS = 53° ($\partial a_g/\partial r$ = 0.25 x 10^{-14} sec^{-2})		
v_α	= 0	No Ranging	389.	0.13	0.22
v_δ	= 0	Ranging	0.01	0.098	0.21
② v_α	= 3	No ranging	1443.	0.14	0.22
v_δ	= 0	Ranging	0.01	0.0039	0.22

Trajectory characteristics: $r = 26 \times 10^6$ km; $v_r = 0$; $\delta = 15°$; σ_β = 1 mm/sec σ ranging = 10 m; sample rate = 1/two days; sample rate = 1/min.

CHAPTER IV.

THE EFFECT OF STATION LOCATION ERRORS AND UNMODELED FORCES ON ESTIMATION ACCURACY

In this chapter we provide a description of the errors introduced to the state estimates by the presence of 1) errors in specifying the positions of the Earth-based tracking stations and 2) small forces acting on the spacecraft and perturbing its motion. The data compression point of view developed in the previous chapter is used and both the situations where doppler data alone are employed and then when ranging data augment the basic doppler observable are considered. The analysis presented here will be accurate for only short arcs of tracking data (perhaps one to two weeks) but will provide insight to that important case and, in addition, assist in the interpretation of the long-arc analysis presented in Chapter VI.

Both the station location errors and unmodeled forces enter the doppler observable equation in relatively simple ways. Reconsider Equation 3.4. The partials of the compressed data (a through f) with respect to the errors to be considered are:

$$\frac{\partial \bar{a}}{\partial r_s} = \begin{bmatrix} 0 \\ \dot{m}\cos\delta \\ 0 \\ 0 \\ -\dot{m}\sin\delta\,\dot{\delta} \\ -\dot{m}\dot{\eta}\cos\delta \end{bmatrix} \quad \frac{\partial \bar{a}}{\partial \lambda} = \begin{bmatrix} 0 \\ 0 \\ \dot{m}r_s\cos\delta \\ 0 \\ 0 \\ -\dot{m}r_s\sin\delta\,\dot{\delta} \end{bmatrix} \quad \frac{\partial \bar{a}}{\partial a_{NG}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad (4.1)$$

where r_s ≜ the tracking station's distance from the Earth's spin axis

λ ≜ the longitude of the tracking station

a_{ng} ≜ the unmodeled acceleration in the Earth-probe direction

Note that the third component of the station location, the distance above the Earth's equatorial plane, and the remaining two components of acceleration orthogonal to the line of sight do not enter this analysis.

The problem becomes: knowing the effect of the model errors on the data observable, compute the relationship between data errors and state estimates.

Inversion of the Doppler Partial Derivative Matrix

In the previous chapter, no explicit advantage was taken of the fact that the original data had been compressed into an observable vector whose dimension equaled that of the state vector. That is, if

$$\delta \bar{a} = A \, \delta \bar{x} + n$$

Instead of forming the state estimate by

$$\delta \hat{\bar{x}} = (A^T W A)^{-1} A^T W \, \delta \bar{a} \qquad (4.2)$$

we can simply write

$$\delta \hat{\bar{x}} = A^{-1} \, \delta \bar{a} \qquad (4.3)$$

since A is an invertible matrix and is given by Equation 3.7.

The simple structure of this matrix allows its analytic inverse to be taken with relative ease

$$A^{-1} = \frac{1}{\omega r_s}\begin{bmatrix} 0 & \dfrac{2r\dot{\gamma}\sin\delta\cos\delta + \frac{\partial d}{\partial \gamma} - 2r\dot{\delta}^2}{\frac{\partial d}{\partial r}\sin\delta} & \dfrac{\frac{\partial d}{\partial \mu} + 2r\dot{\mu}\dot{\delta}\sin\delta\cos\delta}{\frac{\partial d}{\partial r}\cos\delta} & \dfrac{\omega r_s}{\partial d/\partial r} & \dfrac{2r\dot{\delta}}{\frac{\partial d}{\partial r}\sin\delta} & \dfrac{2r\cos\delta\dot{\mu}}{\partial d/\partial r} \\ 0 & -1/\sin\delta & 0 & 0 & 0 & 0 \\ 0 & 0 & -1/\cos\delta & 0 & 0 & 0 \\ \omega r_s & 0 & 0 & 0 & 0 & 0 \\ 0 & \dot{\delta}/\tan\delta\sin\delta & 0 & 0 & -1/\sin\delta & 0 \\ 0 & \dot{\alpha}/\cos\delta & \dfrac{-\dot{\delta}\tan\delta}{\cos\delta} & 0 & 0 & \dfrac{-1}{\cos\delta} \end{bmatrix} \qquad (4.4)$$

$$A^{-1}\,\mathbf{A}\begin{bmatrix} \frac{\partial r}{\partial a} & \frac{\partial r}{\partial b} & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \frac{\partial r}{\partial f} \\ \frac{\partial \delta}{\partial a} & & & \\ \vdots & & & \\ \frac{\partial \dot{\alpha}}{\partial a} & & & \frac{\partial \dot{\alpha}}{\partial f} \end{bmatrix}$$

where the partials $\frac{\partial d}{\partial r}$, $\frac{\partial d}{\partial \delta}$, $\frac{\partial d}{\partial \alpha}$ are given by (3.8).

Rotation to Cartesian Coordinates

Pre-multiplication of (4.1) by (4.4) will yield the desired result of the sensitivity of the state estimates to variations in the three model parameters under consideration. However, in order to interpret the results and propogate forward in an inertial frame, a rotation from the actual trajectory related system being used to an inertial cartesian system is required. This is easily shown to be:

$$\mathbf{A}\begin{bmatrix} r \\ r_\delta \\ r_\alpha \\ v_r \\ v_\delta \\ v_\alpha \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & r & 0 & 0 & 0 & 0 \\ 0 & 0 & r\cos\delta & 0 & 0 & 0 \\ 0 & -r\dot{\delta} & -r\cos^2\delta & 1 & 0 & 0 \\ \dot{\delta} & v_r & 0 & 0 & r & 0 \\ \dot{\alpha}\cos\delta & -\dot{\alpha}r\sin\delta & v_r\cos\delta & 0 & 0 & r\cos\delta \end{bmatrix} \mathbf{A}\begin{bmatrix} r \\ \delta \\ \alpha \\ v_r \\ \dot{\delta} \\ \dot{\alpha} \end{bmatrix}^T \qquad (4.5)$$

Where, for example, r_δ is the position deviation in the direction of increasing declination measured from the nominal trajectory. All

quantities on the left are in an inertial frame, whereas quantities on the right are measured from the actual probe state. For example, Δv_{r_T} is the error in the probe's radial velocity; Δv_{r_I} is the error in the velocity component along the nominal radial direction, which is inertially fixed.

Analytic Sensitivities to Model Parameters

The doppler analysis can now be completed. First, pre multiplying (4.1) by (4.4)

$$\frac{\partial \dot{\bar{r}}}{\partial r_s} = \begin{bmatrix} \frac{2r}{r_s \frac{\partial d}{\partial r}} \left\{ \frac{\dot{\alpha}^2 \sin\delta \cos\delta + \frac{\partial d}{\partial r}\left(\frac{1}{2r}\right) - \frac{\dot{\delta}}{\tan\delta}}{\tan\delta} - \dot{\delta}^2 - \dot{\alpha}^2 \cos^2\delta \right\} \\ \frac{1}{\tan\delta\, r_s} \\ 0 \\ 0 \\ \dot{\delta}/r_s(\tan^{-2}\delta + 1) \\ 2\,\dot{\alpha}/r_s \end{bmatrix} \tag{4.6}$$

$$r_s \frac{\partial \dot{\bar{r}}}{\partial \lambda} \triangleq \frac{\partial \dot{\bar{r}}}{\partial r_\theta} = \begin{bmatrix} \frac{\partial d/\partial \alpha}{\partial d/\partial r}\left(\frac{1}{r_s}\right) \\ 0 \\ -\frac{1}{r_s} \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{4.7}$$

and

$$\frac{\partial \hat{\bar{r}}}{\partial a_{ng}} \begin{bmatrix} 1/\partial d/\partial r \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad (4.8)$$

The slight transformation from $d\lambda$ to $r_s d\lambda$ has been made so that both components of station location error can be expressed in km. Note from (4.6) and (4.7) that the first immediately apparent effect is that an r_s error produces a perturbation in declination, while a longitude error produces a right ascension perturbation in the estimated state. These effects were first demonstrated by Hamilton and Melbourne (Reference 2). In addition, an r_s error introduces a range error whose magnitude is determined by a rather complex interrelationship between elements of the nominal state. This error can be excessively large in certain circumstances, particularly when $\partial d/\partial r$ becomes small or vanishes. Also, there are small velocity perturbations as well.

The state errors introduced by station longitude errors also include a range term which depends strictly upon the ratio of the geocentric acceleration partials with range to those of right ascension. Finally, an unmodeled acceleration error produces an error in range only as shown. Note that all three error sources produce large range errors when $\partial d/\partial r$ is small. Singularities occur in several state components for both r_s and λ errors when the probe declination is zero.

In rotating to the inertial cartesian coordinate system, one further substitution must be made to keep the expressions obtained from getting unwieldy. Define

$$k \triangleq \frac{\dot{\alpha}^2 \sin\delta \cos\delta + \frac{\partial d}{\partial r}\left(\frac{1}{2r}\right) - \dot{\delta}^2/\tan\delta}{\tan\delta} - (\dot{\delta}^2 + \dot{\alpha}^2 \cos^2\delta)$$

Then using the transformation, (4.5), we obtain the sensitivities in the plane-of-sky, inertial, cartesian frame,

$$\frac{\partial \dot{\bar{x}}}{\partial r_s} \triangleq \frac{\partial}{\partial r_s}\begin{bmatrix} \dot{r} \\ r_\delta \\ r_\alpha \\ v_r \\ v_\delta \\ v_\alpha \end{bmatrix} = \begin{bmatrix} 2r/r_s \, \frac{\partial d}{\partial r} \, k \\ \frac{r}{\tan\delta \, r_s} \\ 0 \\ -r\dot{\delta}/\tan\delta r_s \\ \frac{2r\dot{\delta}k}{r_s \, \partial d/\partial r} + \frac{v_r}{\tan\delta r_s} + \frac{r\dot{\delta}}{r_s}(\tan^{-2}\delta + 1) \\ \frac{2r\dot{\alpha}\cos\delta k}{r_s \, \partial d/\partial r} - \frac{r\dot{\alpha}\sin\delta}{\tan\delta r_s} + \frac{2r\cos\delta\dot{\alpha}}{r_s} \end{bmatrix} \qquad (4.9)$$

The longitude sensitivities are much simpler:

$$\frac{\partial \dot{\bar{x}}}{\partial r_\theta} = \begin{bmatrix} \partial d/\partial\alpha \Big/ \frac{\partial d}{\partial r} \, r_s \\ 0 \\ \frac{-r\cos\delta}{r_s} \\ \frac{-r\cos^2\delta \, \dot{\alpha}}{r_s} \\ \left(\frac{\partial d}{\partial\alpha} \Big/ \frac{\partial d}{\partial r} \, r_s\right) \dot{\delta} \\ (\dot{\alpha}\cos\delta)\left(\frac{\partial d}{\partial\alpha} \Big/ \frac{\partial d}{\partial r} \, r_s\right) - \frac{v_r \cos\delta}{r_s} \end{bmatrix} \qquad (4.10)$$

$$\frac{\partial \dot{\bar{x}}}{\partial a_{ng}} = \begin{bmatrix} 1/(\partial d/\partial r) \\ 0 \\ 0 \\ 0 \\ \dot{\delta}/\partial d/\partial r \\ \dot{\alpha}\cos\delta/\partial d/\partial r \end{bmatrix} \qquad (4.11)$$

Together (4.6) through (4.11) provide a rather complete description of the effects of these model errors over short tracking periods with doppler only data in both coordinate frames of major interest. Numerical examples are deferred until ranging data are re-introduced to the problem.

The Effect of Ranging Data

When range data are employed the problem can no longer be worked in the simple analytic fashion just demonstrated. We begin by presenting a simple approach, done with the square-root estimation algorithms discussed in the Appendix, for the computation of the sensitivity of the least-squares estimator to the modeling errors under consideration. In general notation, the data set δa, is related linearly to the state through the familiar data equation,

$$\delta a = A\delta x + n$$

where n is the additive noise. In this case, the A matrix is given by (3.10) and the inverse covariance of the noise, J, is given by (3.9). Thus pre-multiplying through by $J^{\frac{1}{2}}$,

$$J^{\frac{1}{2}}\,\delta a = J^{\frac{1}{2}}A\delta x + J^{\frac{1}{2}}n$$

An orthonormal transformation P that places $J^{\frac{1}{2}}A$ in upper triangular form is found and the indicated estimation procedure as outlined in the Appendix is executed

$$\begin{bmatrix} \hat{\delta x} - \delta x \\ \hline \hat{n}_z \end{bmatrix} = PJ^{\frac{1}{2}}n$$

Thus the upper n x m portion (n, m is the dimension of the state, data vector; in this case 6, 7) of $PJ^{\frac{1}{2}}$ specifies the relationship between data noise and estimation error. Finally, if n is replaced by the m x 3 matrix, $\partial n/\partial r_s$, r_θ, a_{ng}, the sensitivity matrix of errors with respect to the three parameters being studied - station locations and non-gravitational forces - is found.

$$\begin{matrix} 6\{ \\ 1\{ \end{matrix} \begin{bmatrix} \tilde{s}_x \\ \hline s_r \end{bmatrix} \triangleq \begin{bmatrix} \dfrac{\partial \hat{x}}{\partial r_s, r_\theta, a_{ng}} \\ \hline \partial r/\partial r_s, r_\theta, a_{ng} \end{bmatrix}$$

The bottom row gives the change in the residuals, i.e., the component of the data orthogonal to the column space of $J^{\frac{1}{2}}A$, with the parameter variations. This is a single number per parameter and conveniently gives the norm of the total residual vector caused by a unit parameter perturbation. This last comment is made parenthetically but we shall return in Chapter VII to what is now a vague point to formally generate machinery that will allow computation of the increase of the weighted sum of squares of the residuals caused by the presence of unmodeled parameters.

A Numerical Example - Response to Unmodeled Forces

In order to apply the foregoing analysis to a realistic situation a set of Viking '75 trajectories is chosen. There are 13 cases in all spanning the launch - arrival date space currently planned for the mission. All but 4 of these cases uce nominal initial conditions at 30 days from Mars encounter. On two separate trajectories, conditions at encounter minus 10 days and encounter minus 60 days are also analyzed. The reason for the multiplicity of cases is to attempt to determine if there are large variations in the answers obtained over the geometry excursions typical of a single mission. We wish to study the influence of arcs considering 1) doppler only data and 2) doppler plus range data. Table 4.1 identifies the case numbers and lists what will prove to be the single most important variable, geocentric declination.

Table 4.2 summarizes the pertinent results from this study. In accordance with Equation 4.8, the doppler - only results show perturbation in the r coordinate only. For a 10^{-11} km/sec^2 applied force, the results are as shown in the first column. There is little variation over the trajectories inspected from a low of 435 km to a high of 545 km. The reason is clear: the error depends largely on the

$$\frac{\partial a_R}{\partial r}$$

term which is determined by the solar gravitation. Since Mars is at a roughly fixed distance from the Sun, the only variable is the EPS angle (Equation 3.8) and this does not vary greatly over the arrival window.

Table 4.1

Viking '75 Trajectories Selected for Small-Force Study

Case No.	Launch Date-Arrival Date	Time	δ (deg.)
I	September 9, '75 August 8, '76	E - 30 days	11.2
II	August 14, '75 July 24, '76	E - 30 days	14.5
III	August 19, '75 September 2, '76	E - 10 days	.48
IV	August 19, '75 September 2, '76	E - 30 days	5.1
V	August 19, '75 September 2, '76	E - 60 days	11.4
VI	September 3, '75 August 11, '76	E - 30 days	8.8
VII	September 13, '75 August 2, '76	E - 30 days	11.9
VIII	August 16, '75 July 1, '76	E - 30 days	18.5
IX	August 1, '75 June 6, '76	E - 10 days	19.9
X	August 1, '75 June 6, '76	E - 30 days	22.0
XI	August 1, '75 June 6, '76	E - 40 days	23.6
XII	August 14, '75 August 13, '76	E - 30 days	10.1
XIII	August 14, '75 August 3, '76	E - 30 days	12.4

Table 4.2

Results for Viking '75 Small Force Study

Assumed Force of 10^{-11} km/sec^2 in Radial Direction

Case II	Doppler Only Δr	Doppler & Range 11 Days of Tracking $\Delta\delta$ (km)	ΔRA km	$\Delta\dot{\delta}$ m/sec	$\Delta\dot{\alpha}\cos\delta$ m/sec	Residual	δ
I	538	21.3 (18.)	-13.4 (-11.)	-.046 (-.038)	.031 (.025)	.436	11.2°
II	545	15.2 (12.3)	-16 (-13)	-.032	.036	.48	14.5°
III	515	23 (23)	8.6E-3 (8.6E-3)	-.1 (-.1)	.002E-2 (.002E-2)	.1E-1	.478°
IV	535	38.8 (35.9)	-5.7 (-5.3)	-.082 (-.075)	.013 (.012)	.284	5.1°
V	555	21.6 (18.1)	-13.6 (-11.4)	-.047 (-.034)	.032 (.026)	.44	11.4°
VI	545	27.7 (24)	-10.5 (- 9.2)	-.058 (-.052)	.024 (.020)	.386	8.8°
VII	545	19.3 (16.0)	-14.4 (-11.9)	-.042 (-.035)	.032 (.027)	.45	11.9°
VIII	535	8.9 (6.99)	-18.6 (-14.6)	-.0195 (-.025)	.044 (.0346)	.52	18.5°
IX	540	7.7 (5.99)	-19.3 (-15.)	-.0165 (-.0125)	.044 (.0344)	.537	19.9°
X	520	4.95 (3.74)	-20.7 (-15.6)	-.01 (-.0078)	.047 (.037)	.568	22°
XI	435	1.34 (.94)	-22.8 (-16.1)	-.002 (-.0014)	.054 (.039)	.648	23.6°
XII	543	24.4 (20.8)	-12.0 (-10.4)	-.052 (-.045)	.027 (.024)	.414	10.16°
XIII	543	19.3 (16.0)	-14.2 (-11.8)	-.041 (-.034)	.032 (.026)	.455	12.4°

Insertion of ranging data dramatically changes the situation. Note that the 500 km perturbation in position is now reduced to approximately 30-40 km distributed between the right ascension and declination directions. In addition, velocity perturbations arise in the $\dot{\alpha}$, $\dot{\delta}$ directions as shown. These later parameters are variable but the maximum obtained is .1 m/sec in the $\dot{\delta}$ direction for case III.

Thus, if position determination is the main criterion used for evaluating the orbit determination accuracy, adding the range point greatly benefits the accuracy in the vicinity of the data being used. The orbit will not, however, propagate as well because of the appearance of velocity errors. If a gravitational field free propagation model is used as a rough indicator of how these errors will propagate, the .1 m/sec error occuring in case III will propagate to 500 km in approximately 50 days. Of course, it is in a different direction than the original 500 km error, which was along the Earth line of night.

The numbers in the parenthesis are the perturbations arising when the acceleration is included in the estimate. To obtain these numbers, an a priori uncertainty of 10^{-11} km/sec^2 was included in the model. The ability to solve for the acceleration and decrease the resultant error varies from case to case, but is nowhere particularly powerful. The residual column displays the r.s.s. of the weighted residuals after the fit caused by the unmodeled acceleration. Since the expected sum of squares from data noise alone would be unity (this will be shown in Chapter VII), this is not a dramatic increase in any of the cases.

The doppler plus range results are a function of all the nominal

probe coordinates, but a glance at Table 4.2 indicates a strong dependence on geocentric declination. To see this more clearly, Figure 4.1 plots the magnitude (note that the sign between each position and its corresponding velocity coordinate is reversed in all cases) of the perturbations versus this factor. The dependence is marked and qualitatively describes as: for low declination, the perturbations enter the δ and $\dot{\delta}$ directions - as declination increases this is traded for errors in the α and $\dot{\alpha}$ directions. The reason for this behavior is clear. Since

$$\frac{\partial b}{\partial \delta} = -wr_s \sin\delta$$

and

$$\frac{\partial e}{\partial \dot{\delta}} = -wr_s \sin\delta$$

as $\delta \to 0$, δ and $\dot{\delta}$ can be perturbed to remove the d term residual without introducing large changes in the b and e residuals. This is borne out by the residual column of Table 4.2. The topocentric terms involving α and $\dot{\alpha}$ contain no such geometric dependence and are relatively constant from case to case.

In summary, the main conclusions reached here for the short arc are:

1. Small forces can induce relatively large errors in range when using doppler only data.
2. Ranging data removes this error nearly completely, but at the expense of errors in δ, α, $\dot{\delta}$, $\dot{\alpha}$.
3. The position determination is much improved in the vicinity of the data taken when this includes range, but the orbit

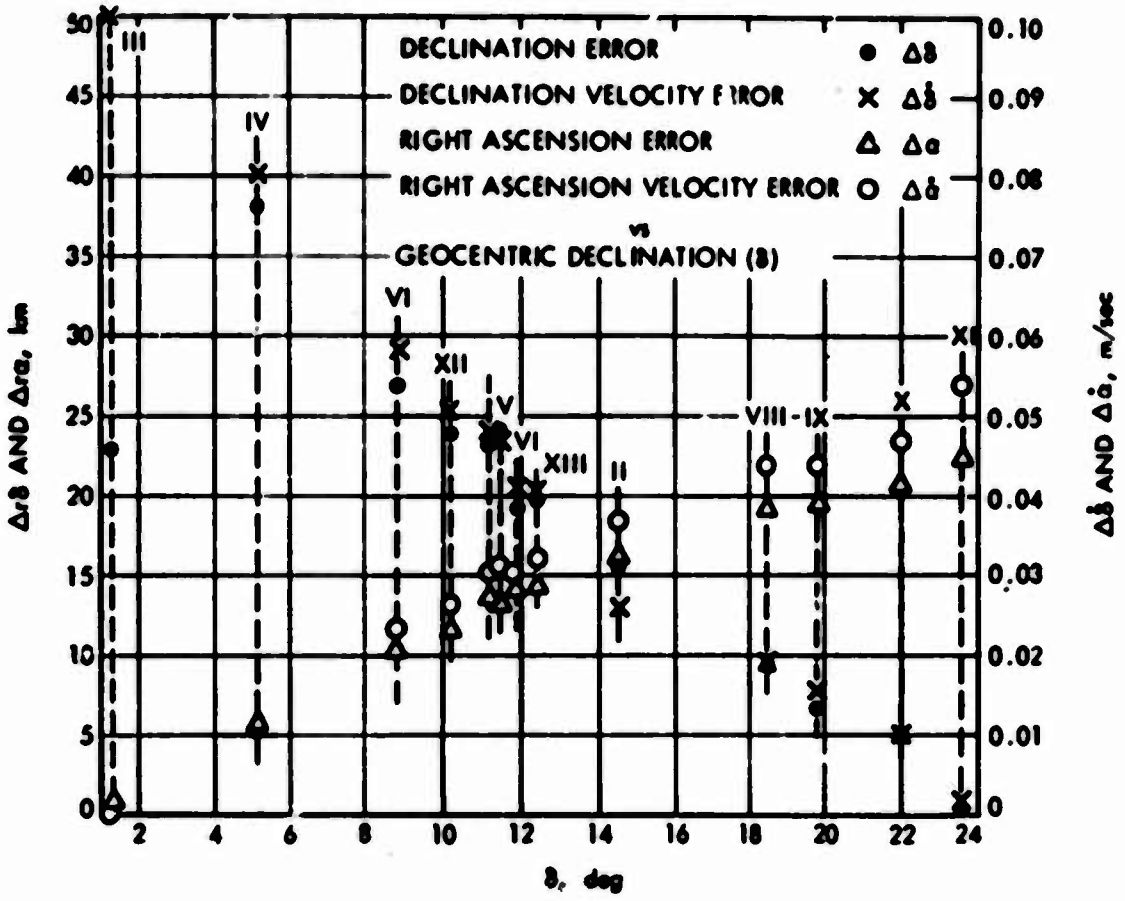


Figure 4.1. Estimation Errors Introduced from a 10^{-11} km/sec^2 Accelerstion.

obtained does not propagate as well. For the Viking cases studied, the break-ever point is on the order of 50 days propagation.

4. The ability to solve for the force over a short time period is modest.

CHAPTER V

A SQUARE-ROOT GENERALIZED CONSIDER OPTION

The preceeding chapter explored the degradation in filter performance due to the existence of two limiting error sources: 1) station location errors, and 2) small accelerations perturbing the motion of the probe. Both developments were only valid for relatively short arcs of tracking data. In this chapter, we wish to develop two techniques for calculating filter performance in imperfectly modeled situations. They will be presented here as self-contained studies and as apparent diversions from the main pursuits of the previous chapter, but will both immediately be applied to the problem of filter response to poorly modeled environments: This time over long tracking data spans. We wish to develop these techniques as an application of the Householder transformation approach to the solution of the linear least squares problem. This technique, which is sometimes referred to as square root filtering, originally came from a method by Householder for inverting square matrices (by rotating them to upper triangular form and inverting with backwards substitution - Reference 9). This was extended to non-square matrices by Businger and Golub (Reference 10). Later, Hanson and Lawson identified two basic algorithms to 1) construct a Householder transformation matrix in compactly stored form and 2) to multiply a vector by such a matrix (Reference 11). Dyer and McReynolds derived sequential formulae for adapting the square root concepts to Kalman-Bucy filtering including the optimal treatment of

process noise (Reference 12).

A short derivation of the use of these techniques for the simplest of estimation problems is presented in Appendix III where it may be compared to the classical least squares and sequential formulae. The following assumes familiarity with that material. We do not deal explicitly with the numerical advantage accrued by using the square-root point of view; the material in References 9-12 is thought to be conclusive on this point. Rather, it is assumed that any filtering or evaluation problem which can be cast in this framework benefits adequately to justify the effort.

The Square Root Consider Option

Dyer and McReynolds (Reference 13) have extended the application of the Householder algorithm to include the case where some of the unknown parameters are not included in the estimation set. It is desired to calculate the contribution of these parameters, referred to as "consider" parameters, to the resultant description of the estimation errors. We recall that development briefly here. Suppose that an information array has been obtained in terms of two classes of parameters, x and y. That is

$$\left[\begin{array}{cc} R_x & R_{xy} \\ 0 & R_y \\ \hline 0 & 0 \end{array}\right]\left[\begin{array}{c} \hat{x} \\ \hat{y} \end{array}\right] + \hat{n} = z - \left[\begin{array}{cc} R_x & R_{xy} \\ 0 & R_y \\ \hline 0 & 0 \end{array}\right]\left[\begin{array}{c} x \\ y \end{array}\right] + PW^{\frac{1}{2}}n \quad (5.1)$$

Suppose that instead of estimating x and y, x only is estimated and y is left at the a priori value of zero. Setting $\hat{y}$ to zero and

pre-multiplying (5.1) by

$$\begin{bmatrix} R_x^{-1} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

results in,

$$\begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{bmatrix} = \begin{bmatrix} x + R_x^{-1}R_{xy}y \\ R_y y \\ 0 \end{bmatrix} + \begin{bmatrix} R_x^{-1} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} W^{\frac{1}{2}} n \qquad (5.2)$$

Transposing the x in (5.2) to the left hand side, post-multiplying the resultant equation by its transpose and taking ensemble averages yields:

$$E\left[(\hat{x} - x)(\hat{x} - x)^T\right] \triangleq \Lambda_x = R_x^{-1}R_x^{-1^T} + R_x^{-1}R_{xy}\Lambda_y R_{xy}^T R_x^{-1^T} \qquad (5.3)$$

where

$$\Lambda_x \triangleq E\left[y\, y^T\right]$$

That is, y is thought of as a random variable (uncorrelated with x) with covariance Λ_y. Whether the resultant estimator of x is unbiased will depend upon the mean of y: this is usually assumed as zero. Equation (5.3) can easily be shown to be equivalent to the classical least squares result for the same circumstances as derived in the Appendix.

Generalizing to Incorrect Apriori Information and Cross Correlation Between the Estimated and Considered Parameters

In both the simple circumstances of estimating all the relevant variables and considering a subset of those, it has been assumed that the apriori covariance used in the filter design for the estimated variables has been correct. In addition, the apriori cross

correlation between the estimated and consider variables has been assumed zero. Other investigators have studied the effects of incorrect a priori on the estimated parameters both from the batch and sequential filter points of view (See Soong, Reference 14, and Nishimura, Reference 15).

It can be useful in several ways to combine the notions of incorrect a priori with the consider option to aid in practical filter design. As a concrete example, suppose that two parameters exist that enter the data in nearly a linearly dependent manner. It would seem reasonable as a practical matter to estimate only one of these variables, but input to the filter an a priori uncertainty high enough to (partly) compensate for not estimating the second variable. In order to accurately evaluate the effects of doing this, we need to reflect that we have used incorrect (too high) a priori uncertainty on the first variable and have not estimated the second. This, and other similar notions, can be treated more rigorously if any apriori cross correlation between the x and y variables can be handled by the derived algorithms.

The exact problem to be investigated is the following: the filter will estimate the x variables and will include apriori data regarding these variables whose uncertainty is specified as

$$E\left[(\hat{x} - x)(\hat{x} - x)^T\right]_{a\ priori} = \Lambda_{x_c}$$

The actual data is influenced by an additional set of variables, y. Moreover, Λ_{x_c} is not necessarily an accurate model of the apriori data. The true situation is

$$\begin{bmatrix} E\left[(\hat{x}-x)(\hat{x}-x)^T\right] & E\left[(\hat{x}-x)(\hat{y}-y)^T\right] \\ E\left[(\hat{y}-y)(\hat{x}-x)^T\right] & E\left[(\hat{y}-y)(\hat{y}-y)^T\right] \end{bmatrix} = \begin{bmatrix} \Lambda_x & \Lambda_{xy} \\ \Lambda_{xy}^T & \Lambda_y \end{bmatrix}$$

a priori

Additional data, relating to both x and y is taken and adjoined to the assumed a priori data.

$$\begin{bmatrix} \hat{x}_{ap} \\ z \end{bmatrix} \triangleq z_t = \begin{bmatrix} I & 0 \\ A & B \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \epsilon_{ap} \\ \epsilon \end{bmatrix} \tag{5.4}$$

where

$\hat{x}_{ap}$ is the apriori estimate of x having dimension n

z is the m dimensional data vector

y is the p dimensional set of consider variables

ϵ_{ap}, ϵ is the error in $\hat{x}_{ap}$ and z respectively

$$E\left[\epsilon_{ap}\right] = 0 \qquad E\left[\epsilon_{ap}\epsilon_{ap}^T\right] = \Lambda_x$$

$$E\left[\epsilon\right] = 0 \qquad E\left[\epsilon\epsilon^T\right] = \Lambda_\epsilon$$

The total data vector, z_t, is weighted by the inverse square-root of the assumed error covariance,

$$\begin{bmatrix} \Lambda_{x_c}^{-\frac{1}{2}} & 0 \\ 0 & \Lambda_\epsilon^{-\frac{1}{2}} \end{bmatrix}$$

yielding,

$$\begin{bmatrix} \Lambda_{x_c}^{-\frac{1}{2}} & 0 \\ 0 & \Lambda_\epsilon^{-\frac{1}{2}} \end{bmatrix} z_t \triangleq z_t' = \begin{bmatrix} \Lambda_{x_c}^{-\frac{1}{2}} & 0 \\ \Lambda_\epsilon^{-\frac{1}{2}}A & \Lambda_\epsilon^{-\frac{1}{2}}B \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \Lambda_{x_c}^{-\frac{1}{2}} & \epsilon_{ap} \\ \Lambda_\epsilon^{-\frac{1}{2}} & \epsilon \end{bmatrix} \tag{5.5}$$

Equation (5.5) is triangularized by an (n+m) x (n+m) orthonormal matrix, P, resulting in

$$Pz_t' \triangleq \begin{bmatrix} z_x'' \\ z'' \end{bmatrix} = \begin{bmatrix} R_x & R_{xy} \\ 0 & R_y \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} P_{11}\Lambda_{x_c}^{-\frac{1}{2}} e_{ap} + P_{12}\Lambda_e^{-\frac{1}{2}} e \\ P_{21}\Lambda_{x_c}^{-\frac{1}{2}} e_{ap} + P_{22}\Lambda_e^{-\frac{1}{2}} e \end{bmatrix} \tag{5.6}$$

where P has been partitioned as

$$\begin{matrix} & \overbrace{}^{n} & \overbrace{}^{m} \\ n & \\ m & \end{matrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$$

Recognizing that the estimate of x is simply

$$\hat{x} = R_x^{-1} z_x'',$$

Then

$$(\hat{x} - x) = R_x^{-1} \left\{ R_{xy} y + P_{11}\Lambda_{x_c}^{-\frac{1}{2}} e_{ap} + P_{12}\Lambda_e^{-\frac{1}{2}} e \right\} \tag{5.7}$$

Multiplying (5.7) by its transpose and taking ensemble averages results in the posteriori covariance on x,

$$\begin{aligned} E\left[(\hat{x} - x)(\hat{x} - x)^T\right] \triangleq \Lambda_x^+ = R_x^{-1} \Big\{ & R_{xy} E\left[yy^T\right] R_{xy}^T + R_{xy} E\left[y\, e_{ap}^T\right] \Lambda_{x_c}^{-\frac{1}{2}T} P_{11}^T \\ & + P_{11}\Lambda_{xc}^{-\frac{1}{2}} E\left[e_{ap} y^T\right] R_{xy}^T + P_{11}\Lambda_{x_c}^{-\frac{1}{2}} E\left[e_{ap} e_{ap}^T\right] \Lambda_{x_c}^{-\frac{1}{2}} P_{11}^T \\ & + P_{12}\Lambda_e^{-\frac{1}{2}} E\left[ee^T\right] \Lambda_e^{-\frac{1}{2}T} P_{12}^T \Big\} R_x^{-1T} \end{aligned} \tag{5.8}$$

Equation (5.8) must be interpreted in terms of the original problem statement to arrive at the final expression. Since $\hat{y}$ is zero,

$$\begin{aligned} E\left[yy^T\right] &= \Lambda_y \\ E\left[y\, e_{ap}^T\right] &= -\Lambda_{xy} \\ E\left[e_{ap} e_{ap}^T\right] &= \Lambda_x \end{aligned}$$

$$E\left[\epsilon\epsilon^T\right] = \Lambda_\epsilon$$

Invoking a consequence of the orthonormality of P, P_{12} can be written in terms of P_{11} as,

$$P_{12}P_{12}^T = I - P_{11}P_{11}^T$$

Substituting the above in (5.8) yields the desired result,

$$\Lambda_x^+ = R_x^{-1}\left\{R_{xy}\Lambda_y R_{xy}^T - R_{xy}\Lambda_{xy}\Lambda_{x_c}^{-\frac{1}{2}T}P_{11}^T - P_{11}\Lambda_{x_c}^{-\frac{1}{2}}\Lambda_{xy}^T R_{xy}^T + P_{11}\left[\Lambda_{x_c}^{-\frac{1}{2}}\Lambda_x\Lambda_{x_c}^{-\frac{1}{2}T} - I\right]P_{11}^T + I\right\}R_x^{-1T} \quad (5.9)$$

Finally, the posteriori covariance between $\hat{x}$ and y is readily seen to be:

$$E\left[(\hat{x} - x)(\hat{y} - y)^T\right] \triangleq \Lambda_{xy}^+ = -R_x^{-1}R_{xy}\Lambda_y + P_{11}\Lambda_{x_c}^{-\frac{1}{2}}\Lambda_{xy} \quad (5.10)$$

The P matrix is not normally explicitly calculated in the standard Householder transformations and special provisions must be made to calculate P_{11}. R Hanson has pointed out in a private communication that the most convenient way of accomplishing this is to adjoin an (n+m)x n matrix to the weighted data vector, $W_{z_t}^{\frac{1}{2}}$, of the following form:

$$P\begin{bmatrix} W_{z_t}^{\frac{1}{2}} \\ \hline I \\ 0 \end{bmatrix} = \begin{bmatrix} z_1' & P_{11} \\ \hline z_2' & P_{21} \end{bmatrix}$$

P_{11} (as well as P_{21}) will appear as a result of this operation. It does, however, require new storage requirements as shown, although if that causes a problem, P_{11} can be formed sequentially and P_{21} never need be explicitly stored.

A Stochastic Square-Root Option

Perhaps the most important advantage of sequential filtering over the classical batch techniques is its ability to deal optimally with stochastic forcing functions in the state dynamics. The simplest of circumstances would be to have a state equation of the form

$$\dot{x} = Fx + w$$

or

$$x_{i+1} = Fx_i + w_i$$

in a discrete-time system, where w (w_i) is a continuous (discrete) white (sequence) process. In this circumstance, the filter does not explicitly estimate w but uses the covariance of w in the matrix Ricatti equation for the covariance of the state estimation errors to make a continuous calculation of the optimum Kalman gains possible. In the case of correlated process noise, an explicit estimate of w may be made by augmenting the state equation to include the time correlated noise. The differential equation governing the correlated noise will be itself driven by white noise (i.e., we are dealing with a linear shaping filter).

It is useful to calculate the effects on filter performance due to unmodeled stochastic processes. That is, the filter is ignorant of the existence of these processes and is optimal with respect to a restricted model only. However, the true filter performance in the presence of these stochastic processes must be calculated.

This procedure might be termed a stochastic consider option. In general, the process will affect both the data and the state. The

development here will consider first the restricted case where the data only is affected (e.g., a random component of station location error) since most of the important principles are illustrated with this simpler model. The technique will then be extended to the more general case. We begin by stating the familiar data equation

$$z_i = Ax + By_i + n_i \tag{5.11}$$

where x is the state being estimated (dimension n)

y_i is the instantaneous value of the consider parameters (dimension p) and the data has been pre-weighted so that the additive noise n_i has zero mean and unity covariance. The transition equation for the y's is given by

$$y_{i+1} = My_i + w_i \tag{5.12}$$

where $||M|| < 1$ to keep the process from growing without bound. The w's have the same dimension as y_i, and are generated as white sequence noise having zero mean. Normally M and the covariance on w_i would be chosen so that the covariance of y_i would remain stationary.

In terms of these quantities, the problem is to evaluate the performance of the filter which estimates x but ignores the influence of the y variables. To this end, suppose that at time t_i just prior to the data taking at that point, an information array is obtained that contains the current estimate of x and the true and computed covariance of $\hat{x}$, y_i, and ε, where ε is an n dimensional, zero mean, random variable of unity covariance whose involvement in the problem will become clear subsequently.

$$\begin{bmatrix} R_x & R_{x\xi} & R_{xy} \\ 0 & I & 0 \\ 0 & 0 & R_y \end{bmatrix} \begin{bmatrix} \hat{x} \\ \xi \\ y_i \end{bmatrix} = \begin{bmatrix} \bar{z}_i \\ \hline 0 \\ \hline 0 \end{bmatrix} \tag{5.13}$$

From (5.13), the estimate is given by

$$\hat{x} = R_x^{-1} \bar{z}_i \tag{5.14}$$

The computed (i.e., the stochastic process free) covariance is

$$\Lambda_{\hat{x}_c} = R_x^{-1} R_x^{-1T} \tag{5.15}$$

and the true, or consider covariance is:

$$\Lambda_{\hat{x}} = R_x^{-1} R_x^{-1T} + R_x^{-1} \begin{bmatrix} R_{x\xi} R_{xy} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & R_y \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ 0 & R_y^T \end{bmatrix}^{-1} \begin{bmatrix} R_{x\xi}^T \\ R_{xy}^T \end{bmatrix} R_x^{-1T} \tag{5.16}$$

Equation (5.16) can be viewed in either of two ways: 1) It arises from the concept of the consider option and is simply an application of (5.3) by recognizing that the

$$\begin{bmatrix} I & 0 \\ 0 & R_y \end{bmatrix}$$

is the inverse square-root on the covariance of the consider parameters. 2) Because of the structure of the right-hand-side of 5.13, even if the consider parameters were estimated, their estimates would be zero. Thus, the left-hand-side of (5.13) can be considered as the inverse square root of the covariance of all the parameters. In that sense (5.16) represents the upper left nxn portion of the full covariance.

Data Processing

Data is taken at t_i in accordance with (5.11), and added to the existing data in the following manner:

$$\begin{bmatrix} R_x & R_{x\xi} & R_{xy} \\ A & 0 & B \end{bmatrix} \begin{bmatrix} \hat{x} \\ \xi \\ y_i \end{bmatrix} \begin{bmatrix} z_i^- \\ \text{----} \\ z_i \end{bmatrix} \qquad (5.17)$$

Equation(5.17) is compacted with a standard Householder transformation resulting in the first n rows as

$$\begin{bmatrix} R'_x , R'_{x\xi} , R'_{xy} \end{bmatrix} \begin{bmatrix} \hat{x} \\ \xi \\ y_i \end{bmatrix} = \begin{bmatrix} z_i^+ \end{bmatrix} \qquad (5.18)$$

This data insertion operation has been performed as if nothing below the first row existed. This is best justified on a simpler problem. Suppose there are but two classes of variables, x and y as in (5.1). That is,

$$\begin{bmatrix} R_x & R_{xy} \\ 0 & R_y \end{bmatrix} \begin{bmatrix} \hat{x} \\ y \end{bmatrix} = \begin{bmatrix} z \\ 0 \end{bmatrix}$$

where $R_y^{-1} R_y^{-1T}$ is the covariance on y. Additional data is taken and adjoined as follows.

$$\begin{bmatrix} R_x & R_{xy} \\ 0 & R_y \\ A & B \end{bmatrix} \begin{bmatrix} \hat{x} \\ y \end{bmatrix} = \begin{bmatrix} z \\ 0 \\ z' \end{bmatrix}$$

Row interchanges can be made without changing the problem. Re-write

the above as

$$\begin{bmatrix} R_x & R_{xy} \\ A & B \\ 0 & R_y \end{bmatrix} \begin{bmatrix} \hat{x} \\ y \end{bmatrix} \qquad \begin{bmatrix} z \\ z' \\ 0 \end{bmatrix}$$

A Householder transformation is applied, resulting in

$$\begin{bmatrix} R_x' & R'_{xy} \\ 0 & R'_y \\ 0 & 0 \end{bmatrix}$$

Notice that R'_x, R'_{xy} obtained in this fashion is identical to what would have been obtained if the transformation had been applied to

$$\begin{bmatrix} R_x & R_{xy} \\ A & B \end{bmatrix}$$

Further, R'_y is not needed. The covariance on x is now

$$(R_x'^{-1})(R'^{-1}_x)^T + R_x'^{-1} R'_{xy} \Lambda_y R'^{T}_{xy} (R'^{-1}_x)^T$$

This completes the justification.

Returning to the full problem, the estimate at this juncture would be

$$\hat{x} = R'^{-1}_x z_i^+$$

and the computed and consider covariance would be calculated as in (5.15) and (5.16). The primed quantities would be used, of course; the covariance on ξ and y_i would remain unchanged.

Mapping

Equation (5.18) must be referred to t_{i+1}. This is done by solving (5.12) for y_i,

$$y_i = M^{-1}y_{i+1} - M^{-1}w_i$$

and re-writing 5.18 in terms of x, y_{i+1}, and w_i. We obtain

$$\begin{bmatrix} R'_x & R'_{x\xi} & R'_{xy} \\ 0 & I & 0 \\ 0 & 0 & R_y \end{bmatrix} \begin{bmatrix} \hat{x} \\ \xi \\ M^{-1}y_{i+1} - M^{-1}w_i \end{bmatrix} \qquad (5.19)$$

Performing the indicated matrix multiplication, effecting certain row interchanges and adding the a priori information on w:

$$\begin{bmatrix} R'_x & R'_{x\xi} & -R'_{xy}M^{-1} & R'_{xy}M^{-1} \\ 0 & I & 0 & 0 \\ 0 & 0 & -R_yM^{-1} & R_yM^{-1} \\ 0 & 0 & \Lambda_w^{-\frac{1}{2}} & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \xi \\ w_i \\ y_{i+1} \end{bmatrix} \qquad (5.20)$$

The last row of 5.20 is to reflect that w_i has mean zero and covariance Λ_w. The most convenient way of expressing this is to add a p dimensional data vector whose a priori value is zero related to w through an identity transformation.* That is

$$z_w = 0 = Iw_i + \epsilon_w$$

where $E[\epsilon_w] = 0$ and $E[\epsilon_w \epsilon_w^T] = \Lambda_w$

The dummy measurements are weighted by pre-multiplying by $\Lambda_w^{-\frac{1}{2}}$ to obtain

$$0 = \Lambda_w^{-\frac{1}{2}} w_i + n_w$$

where $E[n_w n_w^T] = I$

*This particular operation was first used by Dyer and McReynolds in their derivation of an optimal square-root sequential filter for process noise (Reference 12).

which is the last row of (5.20).

If (5.20) is compacted with a Householder transformation, the following array is obtained.

$$\begin{bmatrix} R_x' & R_{x\xi}' & R_{xw} & R_{xy} \\ 0 & I & 0 & 0 \\ 0 & 0 & R_w & R_{wy} \\ 0 & 0 & 0 & R_y \end{bmatrix} \begin{bmatrix} \hat{x} \\ \xi \\ w_i \\ y_{i+1} \end{bmatrix} \tag{5.21}$$

where the 's have been retained to emphasize that the lst 2 rows remain unchanged by this transformation. The lower right 2 x 2 portion represents the inverse square-root of the covariance ow w and y_{i+1}. These are correlated stochastic variables. If the contribution of their effects to the true covariance were calculated at this juncture, it would be

$$R_x^{-1}\begin{bmatrix} R_{xw} & R_{xy} \end{bmatrix}\begin{bmatrix} R_w & R_{wy} \\ 0 & R_y \end{bmatrix}^{-1}\begin{bmatrix} R_w & R_{wy} \\ 0 & R_y \end{bmatrix}^{-1T}\begin{bmatrix} R_{xw}^T \\ R_{xy}^T \end{bmatrix} R_x^{-1\,T}$$

For reasons to be clear presently, it is desired to replace the w_i, y_{i+1}, pair with a new pair, ξ, y_{i+1}, where ξ is a p dimensional random variable independent of y_{i+1} (and ξ) with covariance (for convenience) I. In other words, abandoning the matrix notation for the moment and writing out the first row of (5.21)

$$z_{i+1} = R'_x x + R_{x\xi}\xi + R_{xw} w_i + R_{xy} y_{i+1} + n \tag{5.22}$$

We wish an equivalent data vector of the form

$$z_{i+1} = R'_x x + R'_{x\tilde{\xi}}\tilde{\xi} + R_{\tilde{x}\tilde{\xi}}\tilde{\xi} + R'_{xy} y_{i+1} + n \tag{5.23}$$

For equivalency, two conditions must be observed:

1) the contribution of $\tilde{\xi}$ and y_{i+1}

to the total error covariance from 5.23 must be identical to the contribution of w_i and y_{i+1} from 5.22).

2) the cross correlation of z_{i+1} with y_{i+1}

must be the same in 5.23 and 5.22. This is needed when adding new data which involves y_{i+1}. Since in effect 5.23 plays the role now of apriori data, its relationship to y_{i+1} must be expressed properly.

Invoking the 2nd condition first,

$$R_{xw}\, E\left[w_i y_{i+1}^T\right] + R_{xy}\, E\left[y_{i+1} y_{i+1}^T\right] = R'_{xy}\, E\left[y_{i+1} y_{i+1}^T\right]$$

Recalling that

$$y_{i+1} = My_i + w_i$$

then $E\left[w_i\, y_{i+1}^T\right] = \Lambda_{w_i}$

and $E\left[y_{i+1}\, y_{i+1}^T\right] = R_y^{-1} R_y^{-1T} = \Lambda_y$

(In the usual stationary circumstances, Λ_y will not be a function of i)

Thus,

$$R'_{xy} = R_{xw}\, \Lambda_w\, \Lambda_y^{-1} + R_{xy} \tag{5.24}$$

the first requirement implies that

$$\left[R_{xw}\ R_{xy}\right] \begin{bmatrix} \Lambda_w & \Lambda_w \\ \Lambda_w & \Lambda_y \end{bmatrix} \begin{bmatrix} R_{xw}^T \\ R_{xy}^T \end{bmatrix} \tag{5.25}$$

$$= \left[R_{x\xi},\ R'_{xy}\right] \begin{bmatrix} I & 0 \\ 0 & \Lambda_y \end{bmatrix} \begin{bmatrix} R^T_{x\tilde{\xi}} \\ R^{T'}_{xy} \end{bmatrix}$$

Carrying out the multiplication indicated by 5.25, substituting for

value obtained for R'_{xy} in (5.24), and solving for $R_{x\tilde{\xi}} R_{x\tilde{\xi}}^T$

$$R_{x\tilde{\xi}} R^T_{x\tilde{\xi}} = R_{xw} \Lambda_w R_{xw}^T - R_{xw} \Lambda_w \Lambda_y^{-1} \Lambda_w R_{xw}^T \tag{5.26}$$

$$= R_{xw} \Lambda_w^{\frac{1}{2}} (I - \Lambda_w^{\frac{1}{2}} \Lambda_y^{-1} \Lambda_w^{\frac{1}{2}T}) \Lambda_w^{\frac{1}{2}T} R_{xw}^T$$

Finally since $\Lambda_y \geq \Lambda_w$, 5.25 can be written in explicit square root form as

$$R_{x\tilde{\xi}} = R_{xw} \Lambda_w^{\frac{1}{2}} (I - \Lambda_w^{\frac{1}{2}} \Lambda_y^{-1} \Lambda_w^{\frac{1}{2}T})^{\frac{1}{2}} \tag{5.27}$$

Since ξ and $\tilde{\xi}$ are independent of any of the other variables, their influence can be broken out separately (this of course is the motivation for the transformation to $\hat{\xi}$, y_{i+1} from w_i, y_{i+1}). This contribution is written as

$$R_x^{-1} \left[R_{x\xi}, R_{x\tilde{\xi}} \right] \begin{bmatrix} R^T_{x\xi} \\ R^T_{x\tilde{\xi}} \end{bmatrix} R_x^{-1T} \tag{5.28}$$

Now, if $P P^T = I$, an orthonormal transformation is inserted in the middle of (5.28) and P^T is such that

$$P^T \begin{bmatrix} R^T_{x\xi} \\ R^T_{x\xi} \end{bmatrix} = \begin{bmatrix} \hat{R}_{x\xi}^T \\ 0 \end{bmatrix}$$

Then, the contribution from ξ and $\tilde{\xi}$ can be written in terms of a single random variable again, which shall continue to be denoted as ξ. Finally the post mapping information array is written as

$$\begin{bmatrix} R_x' & \hat{R}_{x\xi} & R'_{xy} \\ 0 & I & 0 \\ 0 & 0 & R_y \end{bmatrix} \begin{bmatrix} x \\ \xi \\ y_{i+1} \end{bmatrix} = \begin{bmatrix} z_{i+1} \\ 0 \\ 0 \end{bmatrix}$$

But equation 5.29 is of the same form as 5.13 except that one full cycle of data taking and mapping has been completed. The reason for the original introduction of the ξ variable should now be clear. It is to express the errors in x that arise from accumulated effects which are orthogonal to the current y. Its dimension is the same as x.

Extension to the General Case

If it is now desired to consider the general case where the process affects not only the data but the state as well, the procedure differs considerably in detail but not in concept from that just described. When employing a sequential estimation that optimally treats process noise, no augmentation of state is necessary since the indirect effect of the process noise on the data through the perturbation of state is automatically accounted for when the covariance matrix of state estimation errors is increased by the process noise during the mapping operation. In the case where we are considering the process noise, the information array, R_x, expresses the information concerning a hypothetical process noise free system. An auxiliary variable is needed to play the role of representing the accumulated effects of the process noise on the current data. Define

$x(0) \triangleq$ initial state of the system at t_o

$x_i \triangleq$ current system state at t_i

$\Delta x_i(0) \triangleq$ current effective departure from the original initial system state due to process noise. $\Delta x_i(o)$ is such that

$$x_i = U(t_i, t_o)(x(o) + \Delta x_i) \tag{5.29}$$

Figure 5.1 illustrates the meaning of Δx_i pictorially. Thus, for

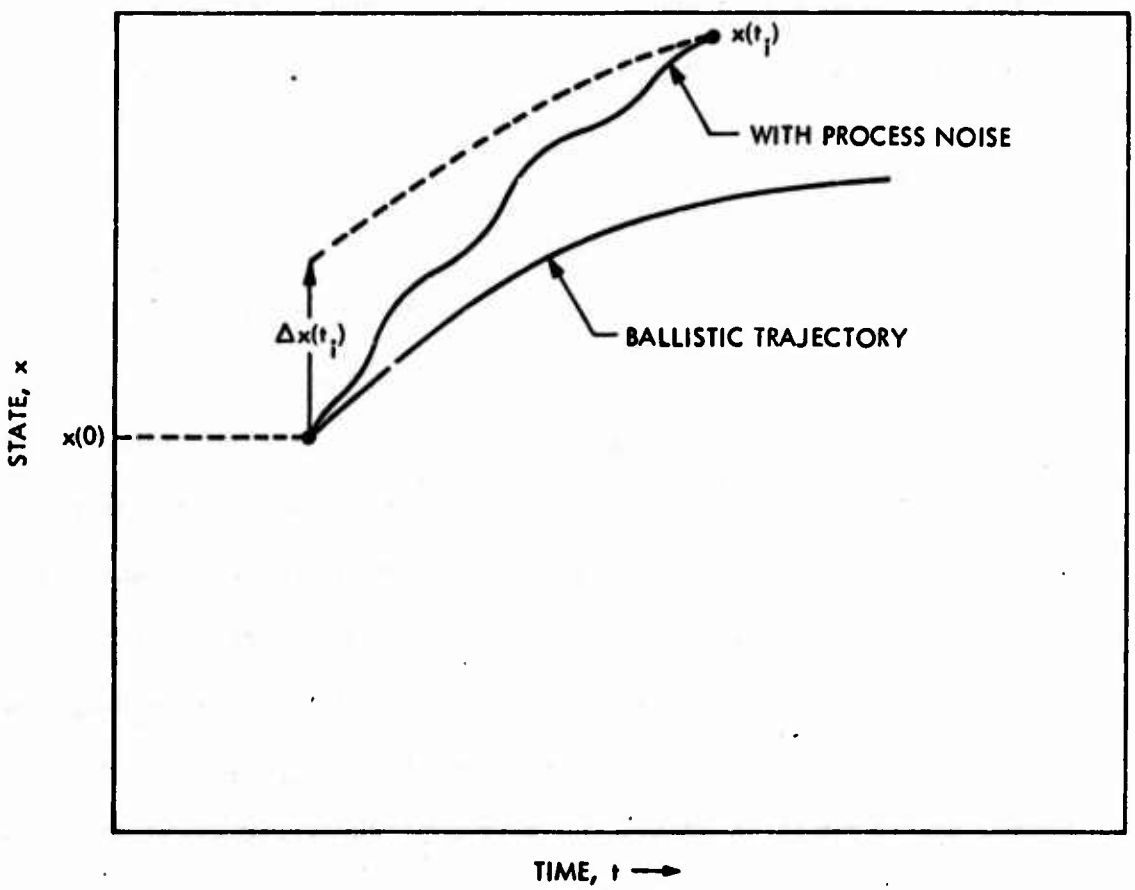


Figure 5.1. Definition of $\Delta x\ (t_i)$.

convenience we shall assume we are estimating not the current state, but the original state x(o) and are considering Δx_i and y_i. From Figure 5.1, if

$$x_{i+1} = U(t_{i+1}, t_i)\, x_i + G_i\, y_i \tag{5.30}$$

then

$$\Delta x_{i+1} = \Delta x_i + U^{-1}(t_{i+1}, t_o)\, G_i y_i$$

Considered alternatively, the Δx_i, y_i jointly now play the role as did y_i alone in the previous development. Their combined state transition would be

$$\begin{bmatrix} \Delta x \\ y \end{bmatrix}_{i+1} = \begin{bmatrix} I & U^{-1}G \\ 0 & M \end{bmatrix} \begin{bmatrix} \Delta x \\ y \end{bmatrix}_i + \begin{bmatrix} 0 \\ I \end{bmatrix} w_i \tag{5.31}$$

The data taken at t_i will relate to the instantaneous state x_i (or through 5.29 the sum of x(0)and Δx_i) and the current process noise,y_i.

$$\begin{aligned} z_i &= A(x(0) + \Delta x_i) + B\, y_i + n_i \\ &= A\, x(0) + A\Delta x_i + B\, y_i + n_i \end{aligned} \tag{5.32}$$

The information array just prior to data taking at t_i (analogous to 5.13) is

$$\begin{bmatrix} R_x & R_{x\xi} & R_{x\Delta x} & R_{xy} \\ 0 & I & 0 & 0 \\ 0 & 0 & R_{\Delta x} & R_{\Delta xy} \\ 0 & 0 & 0 & R_y \\ \hline A & 0 & A & B \end{bmatrix} \begin{bmatrix} x(0) \\ \xi \\ \Delta x_i \\ y_i \end{bmatrix} = \begin{bmatrix} z_i \\ \\ 0 \\ \\ \hline z_i \end{bmatrix} \tag{5.33}$$

where the last row shows the subsequent addition of the data at t_i from 5.32.

The procedure follows that used previously with estimation and covariance computations analogous to 5.14 through 5.16. After data is added at t_i, compaction takes place as before (see 5.17 and 5.18). In order to map, 5.31 is solved for $\begin{bmatrix} \Delta x \\ y \end{bmatrix}_i$, obtaining

$$\begin{bmatrix} \Delta x \\ y \end{bmatrix}_i = \begin{bmatrix} I & -U^{-1}GM^{-1} \\ 0 & M^{-1} \end{bmatrix} \begin{bmatrix} \Delta x \\ y \end{bmatrix}_{i+1} + \begin{bmatrix} U^{-1}\,GM^{-1} \\ -M^{-1} \end{bmatrix} w_i \qquad (5.34)$$

This is substituted in 5.33 after compaction, the same algebraic manipulation takes place, the a priori data on w_i is inserted, and, once again, a forward triangularization is performed. The following array is obtained:

$$\begin{bmatrix} R_x & R_{x\xi} & R_{xw} & R_{x\Delta x} & R_{xy} \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & R_w & R_{w\Delta x} & R_{wy} \\ 0 & 0 & 0 & R_{\Delta x} & R_{\Delta xy} \\ 0 & 0 & 0 & 0 & R_y \end{bmatrix} \begin{bmatrix} x(0) \\ \xi \\ w_i \\ \Delta x_{i+1} \\ y_{i+1} \end{bmatrix} \qquad (5.35)$$

Again an independent random variable $\tilde{w}$ must be introduced to replace w_i. The $\begin{bmatrix} \Delta x \\ y \end{bmatrix}_{i+1}$ play the same role as did y_{i+1} before. Still, the algebra is enough different to present that transformation in some detail.

Define $\begin{bmatrix} \Delta x \\ y \end{bmatrix}_{i+1} \triangleq \zeta$

and $\left[R_{x\Delta x}, R_{xy} \right] \triangleq R_{x\Omega}$

We wish to write

$$z = R_x x + R_{x\xi} \xi + R_{xw} w + R_{x\Omega} \Omega \tag{5.36}$$

as an equivalent

$$z = R_x x + R_{x\xi} \xi + R_{x\tilde{\xi}} \tilde{\xi} + R'_{x\Omega} \Omega \tag{5.37}$$

For equivalency it is required that

$$R_{xw} E\left[w \Omega^T \right] + R_{x\Omega} E\left[\Omega \Omega^T \right] = R'_{x\Omega} E\left[\Omega \Omega^T \right]$$

Defining

$$E\left[w \Omega^T \right] \triangleq \Lambda_{w\Omega} \tag{5.38}$$

$$E\left[\Omega \Omega^T \right] \triangleq \Lambda_{\Omega}$$

We obtain

$$R'_{x\Omega} = R_{xw} \Lambda_{w\Omega} \Lambda_{\Omega}^{-1} + R_{x\Omega} \tag{5.39}$$

We now must select $R_{x\tilde{\xi}}$ so that $\tilde{\xi}$ and Ω will contribute as much to the variance in (5.37) as do w and Ω in (5.36). This is done just as in (5.26) through (5.27). In this case the result is

$$R_{x\tilde{\xi}} R_{x\tilde{\xi}}^T = R_{xw} \Lambda_w R_{xw}^T - R_{xw} \Lambda_{w\Omega} \Lambda_{\Omega}^{-1} \Lambda_{w\Omega}^T R_{xw}^T \tag{5.40}$$

It is easily shown that

$$\Lambda_{w\Omega} = \begin{bmatrix} 0 & \Lambda_w \end{bmatrix}$$

and

$$\Lambda_{\Omega}^{-1} = \begin{bmatrix} R_{\Delta x}^T R_{\Delta x} & R_{\Delta x}^T R_{\ell xy} \\ R_{\Delta xy}^T R_{\Delta x} & R_{\Delta xy}^T R_{\Delta xy} + R_y^T R_y \end{bmatrix}$$

Finally

$$R_{x\tilde{\xi}} = R_{xw}\,\Lambda_w^{\frac{1}{2}}\left\{ I - (\Lambda_w^{\frac{1}{2}T})\left[R^T_{\Delta xy}\,R_{\Delta xy} + R^T_y\,R_y\right]\Lambda_w^{\frac{1}{2}}\right\}^{\frac{1}{2}} \qquad (5.41)$$

and as before, $R_{x\xi}$ and $R_{x\tilde{\xi}}$ are combined completing the generalized procedure.

Propagating Forward and Evaluation

The square-root of the information matrix (5.33) can be referred to a new epoch at any time by simply post multiplying by the inverse state transition matrix, i.e.,

$$\begin{bmatrix} R_x U^{-1} & R_{x\xi} & R_{x\Delta x}U^{-1} & R_{xy} \\ 0 & I & 0 & 0 \\ 0 & 0 & R_{\Delta x}U^{-1} & R_{\Delta xp} \\ 0 & 0 & 0 & R_p \end{bmatrix}\begin{bmatrix} x(j) \\ \xi \\ \Delta x_i(j) \\ p \end{bmatrix} = \begin{bmatrix} z \\ 0 \\ 0 \\ 0 \\ + \end{bmatrix}$$

where

$$U \triangleq U(t_j, t_o)$$

and

$$\Delta x_i(j) = U(t_j, t_o)\,\Delta x_i$$

In evaluating for the covariance of the error in the estimate, it must be remembered that the error in the estimate of x(0) is the error of the original injection conditions. It is usually desired to know the error in the effective conditions, i.e.,

$$x - \hat{x} + (\Delta x - \Delta\hat{x})$$

Since $\Delta\hat{x}$ is always zero, we equivalently wish to calculate the

$$E\left[\left\{(x-\hat{x}) + \Delta x\right\}\left\{(x-\hat{x}) + \Delta x\right\}^T\right]$$

It is easily shown that this is

$$R_x^{-1} R_x^{-1T} + R_x^{-1} R_{x\Delta x} \Lambda_{\Delta x} R_{x\Delta x} R_x^{-1T} + R_x^{-1} R_{x\Delta x} \Lambda_{\Delta x}$$

$$+ \Lambda_{\Delta x} R_{x\Delta x}^T R_x^{-1T}$$

where

$$\Lambda_{\Delta x} = R_{\Delta x}^{-1} (R_{\Delta x}^{-1})^T$$

and is the covariance of perturbations due to the process noise.

It may appear somewhat unorthodox to have cast this problem in terms of estimating the initial state, particularly in a problem that seems most naturally to lend itself to estimating the current state in the usual sequential manner. This has been done for two reasons:

1. The algorithms so derived will mate more easily with already existing computational machinery for their application in Chapter VI.

2. It should be clear that with minor modifications (all simplifications) to what was presented, we could have produced an algorithm for optimally estimating $x(0)$ and Δx_i. That is, a square-root, one-pass, forward smoother in discrete form would have resulted. Forward smoothers are discussed by Nishimura, Ref. 16.

A Numerical Example

To clarify the procedure and to demonstrate its validity, a simple numerical example was constructed, one where the proper answers could be calculated by appeal to a much simpler point of view. Suppose both x, w and hence Δx have dimension one. Further, let $y = w$, i.e, there really is no need to deal with the auxiliary variable, y. Finally let the measurements directly measure x and Δx:

$$z_i = x + \Delta x_i + n_i$$

where

$$\Delta x_i = \Delta x_{i-1} + w_{i-1}$$

and

$$E[n_i] = 0, \quad E[n_i\, n_j] = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

$$E[\Delta x_i] = 0, \quad E[\Delta x_i^2] = 1$$

$$E[w_i] = 0, \quad E[w_i\, w_j] = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

A sequence of measurements is obtained:

$$z_1 = x + \Delta x_1 + n_1.$$

$$z_2 = x + \Delta x_2 + n_2 = x + \Delta x_1 + w_1 + n_2$$

$$z_3 = x + \Delta x_1 + w_1 + w_2 + n_3$$

$$\vdots$$

$$z_n = x + \Delta x_1 + \sum_{i=1}^{n-1} w_i + n_n$$

The filter is aware only of the presence of x and n. Thus it will have the form:

$$\hat{x} = \frac{1}{N} \sum_{i=1}^{N} z_i$$

The true performance can be calculated by considering

$$\Delta x_1 + \sum_{i-1}^{N-1} w_i + n_n$$

to be the true measurement noise and calculate the actual filter performance which equally weights the measurements. It is easy to show that for an equal weighting filter where the actual data noise is Λ_e, the true covariance of estimation error is

$$\Lambda_x = (A^T A)^{-1} A^T \Lambda_e A (A^T A)^{-1}$$

where A is the data partial matrix with respect to the estimated parameters. In this case, A is simply

$$A = \begin{bmatrix} 1 \\ 1 \\ 1 \\ . \\ . \\ . \\ 1 \end{bmatrix}$$

So that

$$\Lambda_x = \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_{ij}/N^2$$

where

λ_{ij} is the i, j element of Λ_e

For example, when N = 5,

$$\Lambda_e = \begin{bmatrix} 2 & 1 & 1 & 1 & 1 \\ 1 & 3 & 2 & 2 & 2 \\ 1 & 2 & 4 & 3 & 3 \\ 1 & 2 & 3 & 5 & 4 \\ 1 & 2 & 3 & 4 & 6 \end{bmatrix}$$

and

$$\Lambda_x = \frac{60}{25} = 2.4$$

This procedure has obvious limitations but serves to generate a check for the stochastic consider option just developed. Table 5.1 displays results from each method for the first few data points. It is seen that agreement is maintained to 8 place accuracy, the printout resolution used (computation was performed in double precision).

Table 5.1

Stochastic Consider Option Check

No. of Data Points	Analytical Result	Stochastic Consider Option Result
1	2	2.0000000
2	7/4 = 1.75	1.7500000
3	17/9 = 1.8888888..	1.8888889
4	34/16 = 2.125	2.1250000
5	60/25 = 2.4	2.4000000

CHAPTER VI.

THE LONG-TERM INFLUENCE OF POORLY MODELED FORCES AND STATION LOCATION MOVEMENT ON ESTIMATION ACCURACY

In Chapter IV, the short-term analytic doppler and range tracking models were applied to the problem of determining the influence of tracking station location errors and non-gravitational forces on state estimation accuracy. In particular, the non-gravitational force problem was explored quantitatively for 1) several trajectories, 2) inclusion of range data or not, and 3) attempted solution for the unknown force or not. The thread of that investigation is picked up here and both the station location problem and the small force problem are explored quantitatively, this time with more precise modeling and over long tracking intervals.

In this Chapter, we shall try to take an operational point of view and address problems inherent in using tracking data during a real-time mission when the objectives are to best determine the state and predict the encounter conditions of the trajectory relative to the target planet. This is a problem that is distinct from the post-flight data reduction problem which emphasizes careful solutions of astrodynamical constants, tracking station parameter calibrations, and perhaps refined trajectory estimates.

Missions up through Mariner '69 have relied upon very simple filters in the sense that the parameter list is abbreviated, usually only including the six components of trajectory state and solar

pressure. Tracking data from past missions is used to produce the most careful calibration possible as to the station locations and the important astrodynamical constants that affect the solution using the abbreviated filter. During the flight, Universal Time and Polar Motion which are essential to locate the instantaneous position of the stations relative to an assumed known inertial coordinate frame, are monitored closely. Depending on the accuracy required, the troposphere, ionosphere, and space plasma - all factors directly affecting the data -are calibrated in real-time by means external to the orbit determination process. Finally conventional wisdom has it that, for example in predicting encounter conditions, only a relatively short arc of data should be used in the final estimation process. This data arc should not exceed 30 days and would normally be the last data received.

The reasons for the abbreviated solution vector and the short tracking arc used can be illuminated here but briefly (for a more detailed discussion of these factors and principles as applied to planetary cruise orbit determination, see Reference 17):

1. Station location errors, even carefully calibrated as they are, still are the predominant error source.* Further it is easy to demonstrate that by using 2 or 3 months of data prior to encounter and estimating the station coordinates produces a claimed orbit determination performance substantially superior to that of using the shorter tracking arc and considering the station errors at their apriori values. Why not then follow this apparently more optimal procedure? The first flaw is that this same long arc analysis will also reveal that the station longitudes will be determined to better than .5 m, standard deviation. This is not reasonable. We have been using nearly a decade of interplanetary tracking data and have never obtained

*Currently the DSN tracking station locations are known approximately to 1.5 m(σ_r) in distance off the spin axis and 3 m(σ_{r_λ}) in the longitude direction.

consistency of answers much better than the claimed knowledge of 3 m. It is not likely that an additional 3 months or so of tracking will markedly improve this situation. Therefore, the long arc analysis results are optimistic, and it is not clear that the actual result would in fact not be worse than the simpler short arc result.

2. The small unmodeled forces may be producing actual filter divergence when too long an arc of data is used. Filter divergence is often characteristic of mismodeling of dynamic parameters, but until recently, there has been little hard evidence as to whether this is the case for interplanetary orbit determinations over the time span we are interested in. In the face of this paucity of evidence, divergence has remained a vague fear and prudence has dictated that short tracking arcs be used. Chapter IV has concluded that, because the data employed are almost directly affected by accelerations, unmodeled forces immediately produce erroneous results at the levels determined, but this analysis gives no clue as the their effects in the longer term. Recently J. F. Jordan et al, produced results which indicated that divergence did not occur over several months tracking intervals when constant forces were simply ignored in the filter itself (Reference 18).

We hope to shed some light on both station location and unmodeled force effects here as an application of the computational machinery developed in Chapter V.

Station Location Error Study

The reason, of course, that the formal statistics for a long data arc including station location solutions are optimistic is because of too simplistic a model. In solving for station parameters, the difference between the actual station location and that used in the abstract model is assumed to be a constant vector in the Earth's rotating co-ordinate frame. In actuality, this difference, and hence the station location error, has a significant stochastic component. This component arises from temporal variations in the

i. Universal Time calibration error
ii. Polar Motion calibration error.

In addition, uncalibrated tidal effects may be producing significant station location errors. Finally, the uncalibrated portion of the ionosphere and space plasma tend to produce unmodeled diurnal signatures in the data that are effectively station location shifts.

There is little hope that these effects can be modeled as Markovian stochastic processes accurately enough to include them as a formal part of the filter model. It should be possible to postulate reasonable enough models for these processes to be used as consider variables however. For example, suppose we wish to test two rather simple filters:

1. Estimate state only
2. Estimate state and a constant station location error.

These two filters can be evaluated in the presence of both the fixed and stochastic station errors for a range of possible descriptions that is thought reasonable for the stochastic portion. If filter 2 outperforms filter 1 over this whole range, confidence will be gained that it truly is superior. Moreover, at the same time statistics would be generated for filter 2 that could be believed.

In short, the maxim operating here is: The filter evaluation should include a model that is at least one step (in some sense) more sophisticated than the filter itself.

In using filter 1), that step has been the consider option including the static station location errors. The results can be believed and probably would not be much worse (they might be better) if the refined stochastic point of view were utilized. In using filter 2), a stochastic portion must be postulated if the statistical

results are not to be hopelessly optimistic.

Before proceeding to discuss the results, it should be emphasized that while every attempt has been made to use realistic numbers, a formal and authoritative model of the stochastic station location components is beyond the scope of this work. We are pursuing a point of view on technique here and the results can at best be claimed as representative. Performance of the actual studies which bring change to current operational philosophy are not best carried out in the relative isolation present when pursuing independent research.

To obtain the desired numerical results, a Univac 1108 computer program was designed and built that executed the Generalized Consider algorithms developed in Chapter V. Since it was desired to concentrate on the estimation aspects of the problem, no trajectory or partial derivative packages were built. Rather, the program was designed in such a way that it could be driven by JPL's Double Precision Orbit Determination Program (Ref. 19) and simply use the trajectory information, state transition matrices, and data partials generated there. In this way, we become heir to all the sophisticated modeling and generality built into the DPODP with no effort on our part.

The point of view taken on the stochastic portion of the station locations (and in the following sub-section - on the random forces) was that they were piecewise constant over an interval Δt. Then from interval to interval they were correlated as

$$e^{-\Delta t/\tau}$$

As long as $\Delta t \ll \tau$, this is thought a reasonable approximation of a continuous exponentially correlated stochastic process. With this

approximation, the data partials could be processed in batches with Δt time span and greatly reduce the computations needed compared to that required if all of the steps of the stochastic consider option were executed each data point. All of the station location results presented here assume a correlation time, τ, of 25 days and use a Δt interval of 4 days. Test results using lower Δt's yielded results only negligibly different from those shown here.

The Specific Case

Again a Viking '75 Mars trajectory was selected but this time, the simulated data commenced six months prior to encounter so that arbitrarily long arcs of data could be investigated. One pass of doppler only data from a single station each day was used at an effective data noise of 1 mm/sec for a 1 minute sample. The station was assumed to have station location errors as follows:

	Radius off spin axis	Longitude
Constant	σ_{r_s} = 1.5 m	σ_{r_λ} = 3.0 m
Stochastic	σ_{r_s} = .75 m τ = 25 days	σ_{r_λ} = .75 m τ = 25 days
All cross correlations were taken to be zero		

Figures 6.1-6.3 display the performance of the filter which estimates the 6 components of state only. Shown are the uncertainties of the three components of initial position in the plane-of-sky coordinate system used in Chapters III and IV:

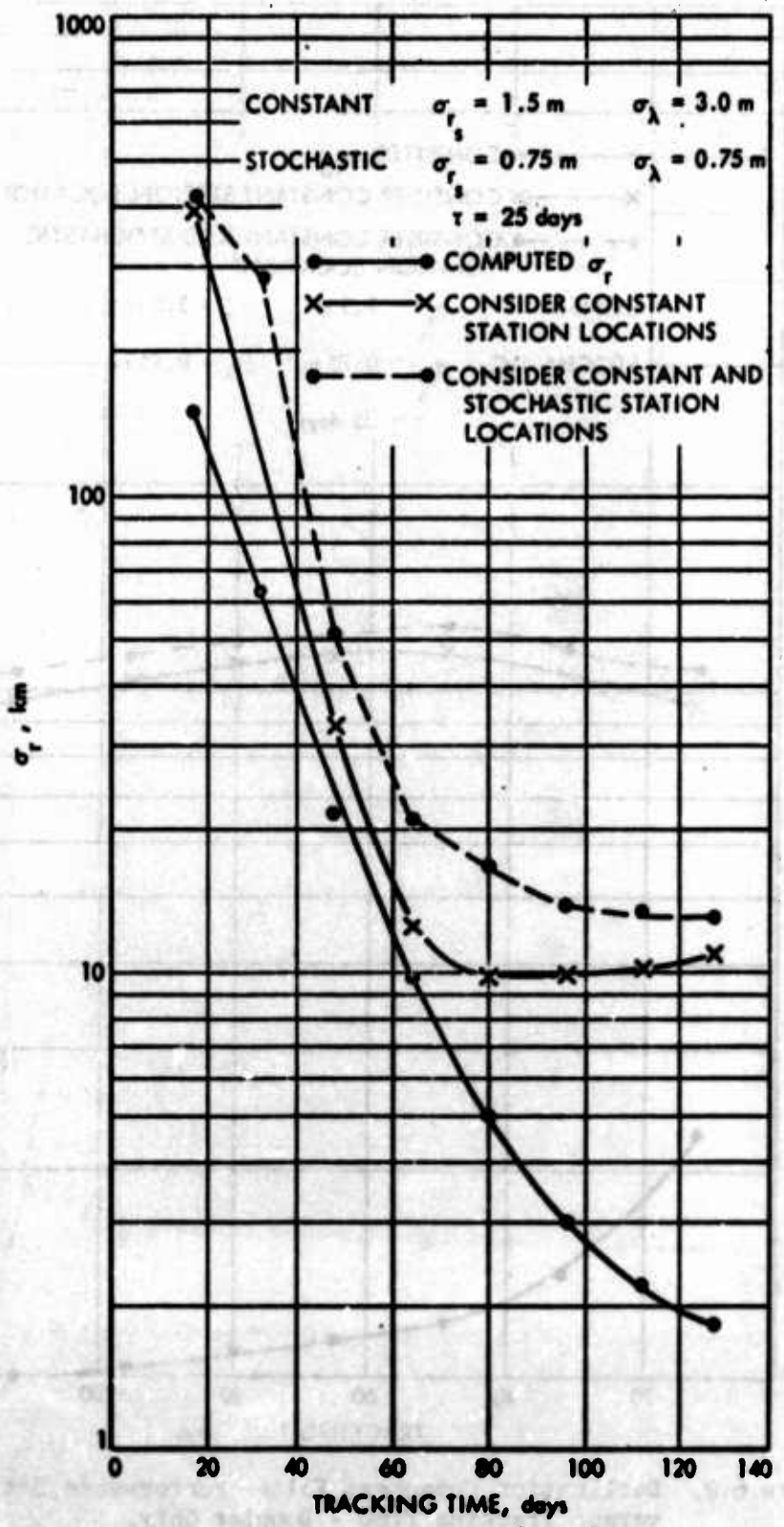


Figure 6.1 Range Component Filter Performance State Only Versus Tracking Time - Doppler Only.

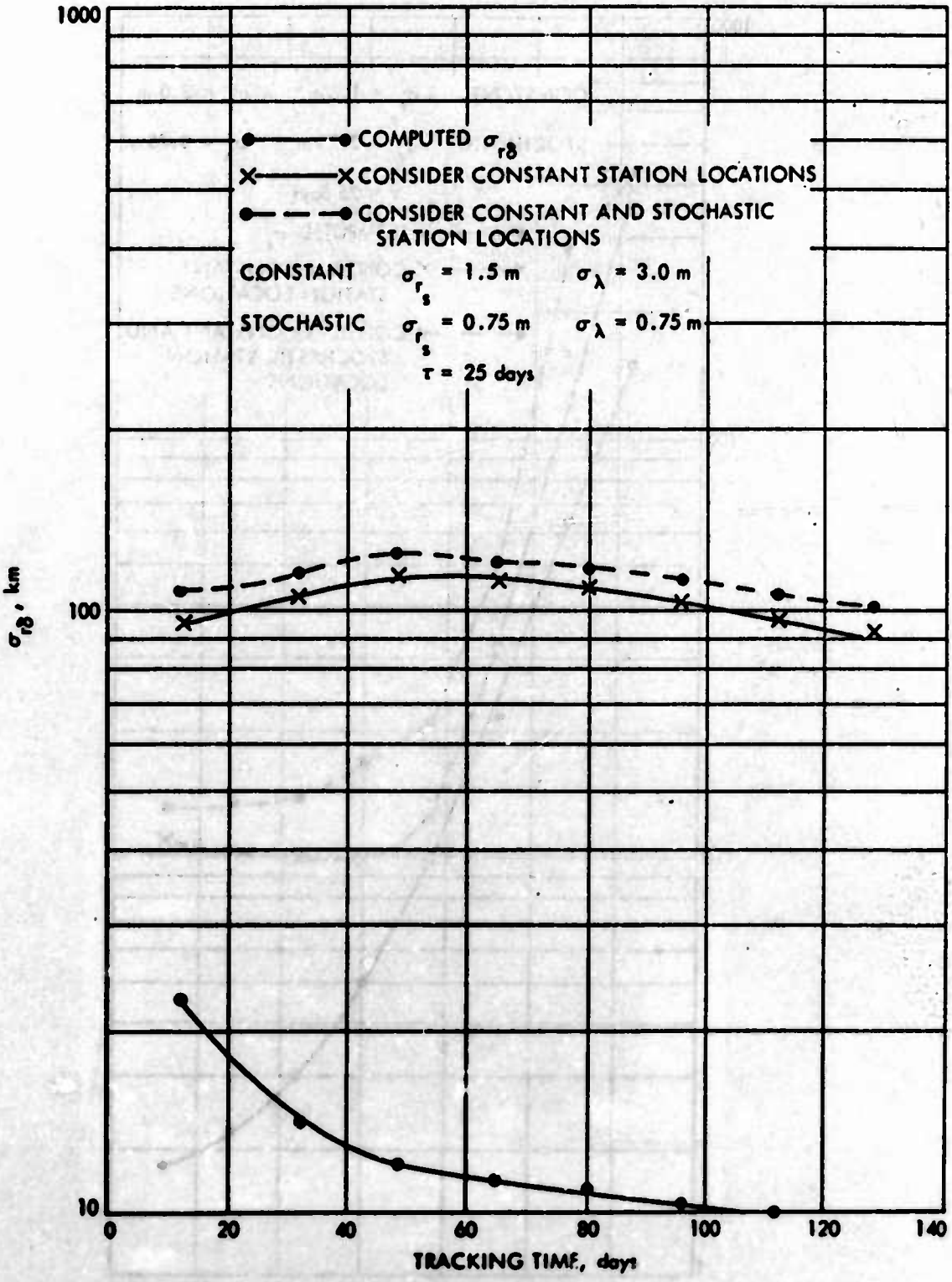


Figure 6.2. Declination Component Filter Performance State Only versus Tracking Time - Doppler Only.

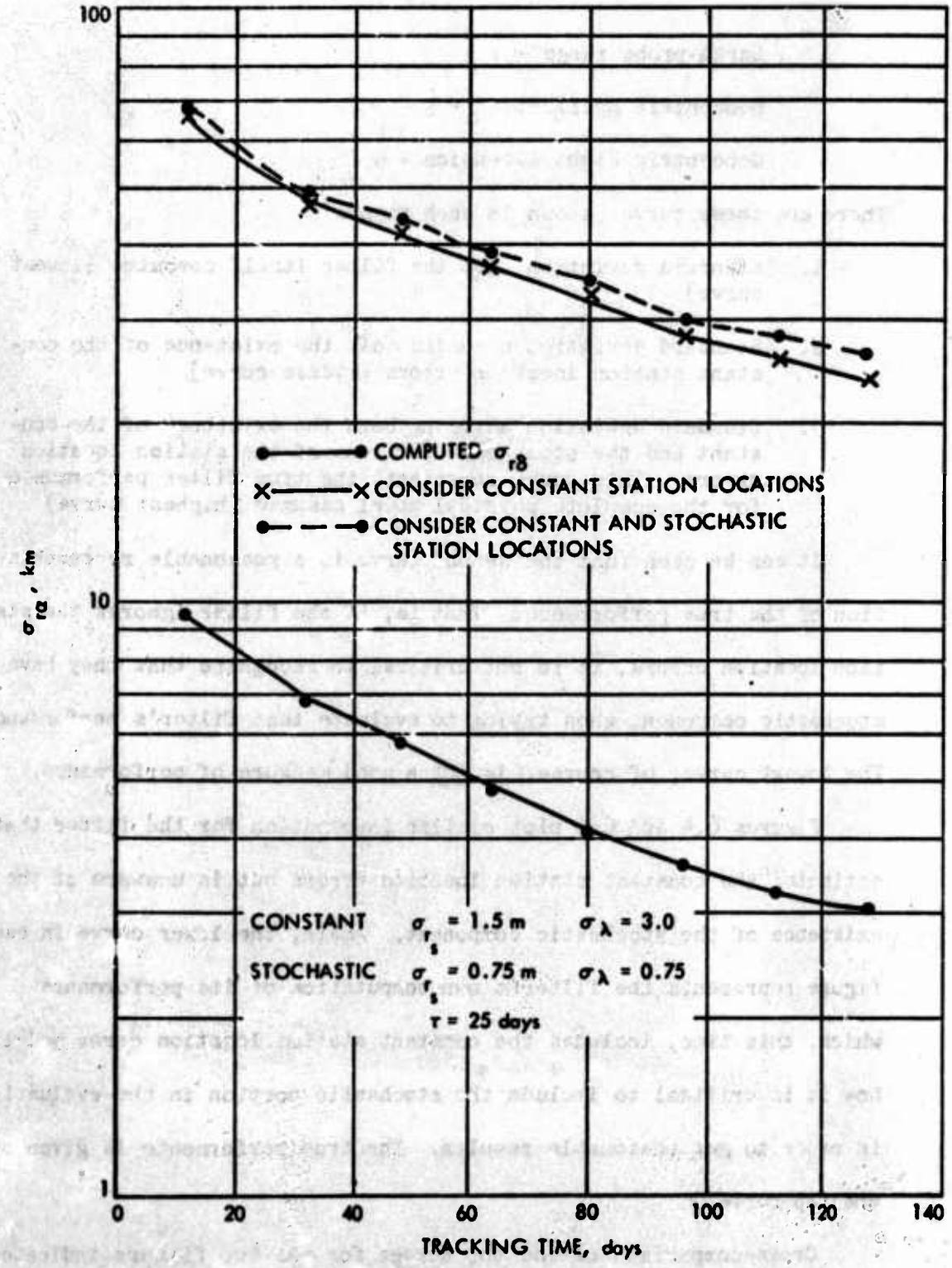


Figure 6.3. Right Ascension Components Filter Performance State Only versus Tracking Time - Doppler Only.

Earth-probe range - r

Geocentric Declination - δ

Geocentric Right Ascension - α

There are three curves shown in each figure:

1. Standard deviation that the filter itself computes (lowest curve)

2. Standard deviation assuming only the existence of the constant station location errors (middle curve)

3. Standard deviation assuming both the existence of the constant and the stochastic portions of the station location errors. This curve represents the true filter performance for the complete physical model assumed (highest curve)

It can be seen that the second curve is a reasonable representation of the true performance. That is, if the filter ignores the station location errors, it is not critical to recognize that they have a stochastic component when trying to evaluate that filter's performance. The lowest curve, of course, is <u>not</u> a good measure of performance.

Figures 6.4 and 6.5 plot similar information for the filter that estimates the constant station location errors but is unaware of the existence of the stochastic component. Again, the lower curve in each figure represents the filter's own computation of its performance which, this time, includes the constant station location error model. Now it is critical to include the stochastic portion in the evaluation in order to get reasonable results. The true performance is given by the top curve.

Cross-comparison of the top curves for the two filters indicates that estimating the station location errors does refine the results; it does not however, revolutionize them. Table 6.1 facilitates this

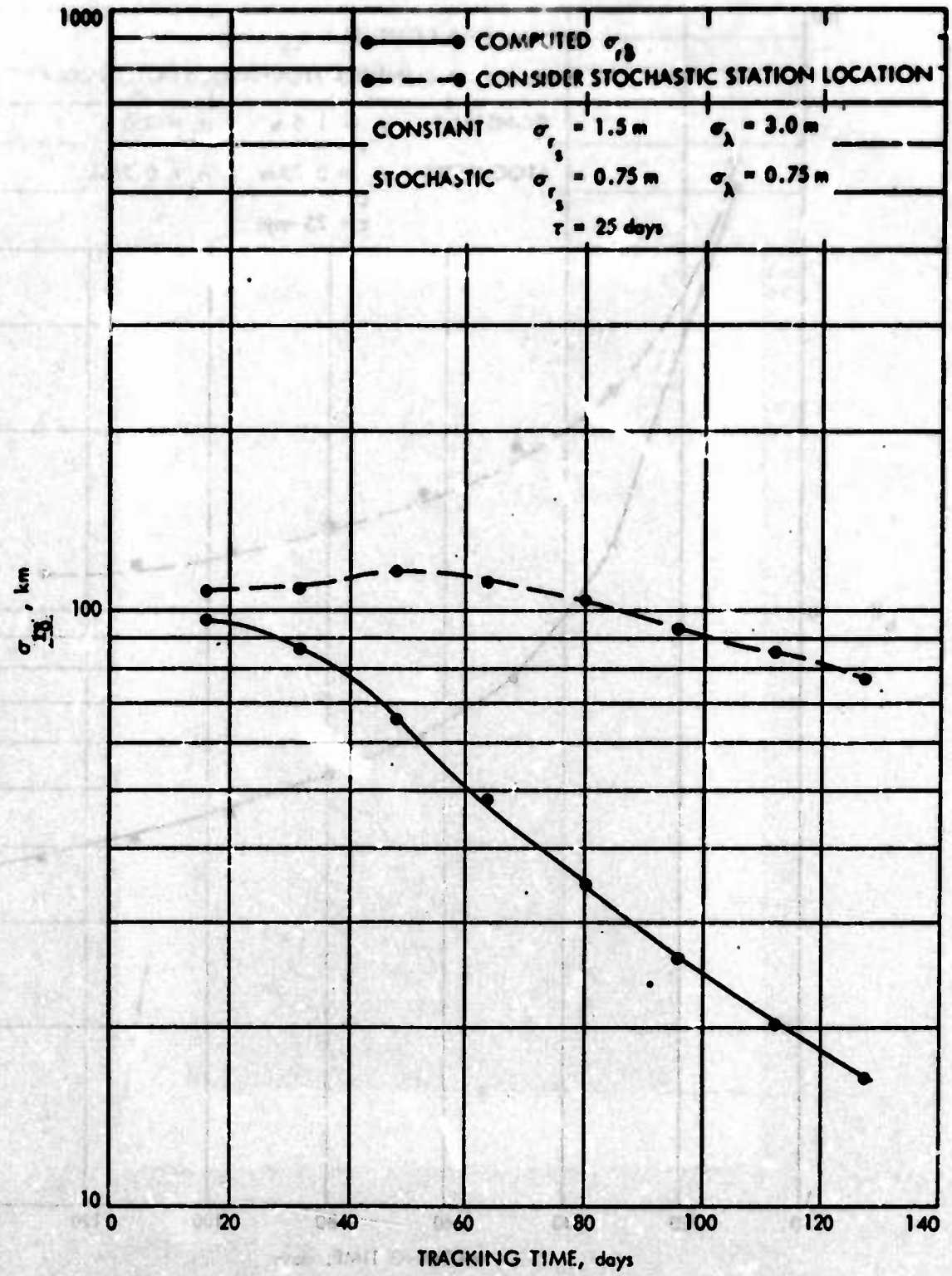


Figure 6.4. Declination Component Filter Performance State and Station Location Solution versus Tracking Time - Doppler Only.

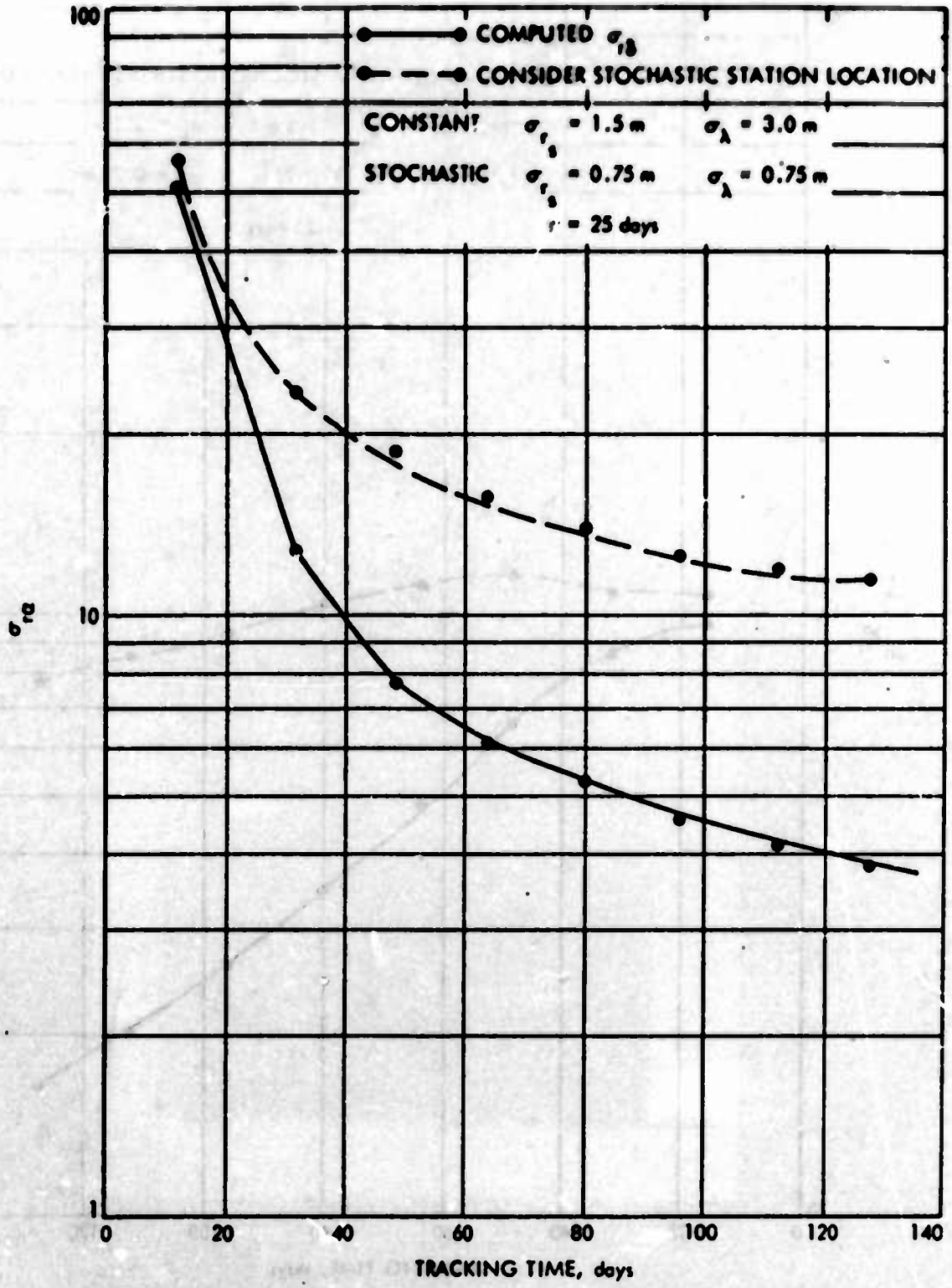


Figure 6.5. Right Ascension Component Filter Performance State and Station Location Solution versus Tracking Time - Doppler Only.

Table 6.1

Comparison of Filter Performance

after 128 Days of Tracking

	σ_r km	$\sigma_{r\delta}$ km	$\sigma_{r\alpha}$ km	$\sigma_{\dot{r}}$ km/sec	$\sigma_{\dot{\delta}}$ km/sec	$\sigma_{\dot{\alpha}}$ km/sec
State Only	13.1	102.8	26.7	2.46×10^{-6}	2.88×10^{-5}	1.78×10^{-6}
State Plus Constant Station Solution	10.1	78.0	11.4	1.18×10^{-6}	1.38×10^{-5}	$7.5? \times 10^{-7}$

Physical Model:

Constant Station Location Uncertainty $\sigma_{r\delta} = 1.5m$ $\sigma_{r\lambda} = 3.0m$

Stochastic Station Location Uncertainties $\sigma_{r\delta} = .75m$ $\sigma_{r\lambda} = .75m$ $\tau = 25$ days

cross-comparison by listing the actual standard deviations for all six of the co-ordinates after 128 days of tracking data. Although each co-ordinate has been improved by extending the parameter set, it is nowhere near as dramatic as if the stochastic portion were not present. Significantly absent is a large improvement in the declination direction, which is always a major contributor to overall performance errors. The per cent improvement in α and the three velocities is more significant - better than a factor of two in each case.

Finally, we wish to present some data concerning the accuracy and validity region of the analytic model discussed in Chapter IV. This is found in Table II which compares the effect of the constant station location errors computed analytically with those obtained from the DFODP whose model assumes little else than Newton and Einstein were correct. Comparing the "total rms" line with the 8 day result, remarkable agreement is found in all but the r position and declination velocity components. Even here, the agreement is quite respectable. As more tracking data is taken (and assumptions in the analytic model are violated), the agreement deteriorates as would be expected. Of particular note is that right ascension position and velocity estimates become less sensitive to station location errors as more data is included. The increase in the range error with time could probably be removed with the addition of ranging data, although this must be carefully done lest an even greater error appear in some other co-ordinate.

Small-Force Study

A study of non-gravitational forces such as presented in

Table 6.2

Comparison of Chapter IV Analytic Model with Exact Model

		r km	A km	r km	r km/sec	A km/sec	O km/sec
ANALYTIC	Δr_s = 1.5m	34.2	93.5	0	$.525 \times 10^{-6}$	1.22×10^{-5}	1.02×10^{-6}
	Δr_j = 3.0m	202.	0	72.0	$.605 \times 10^{-5}$	2.12×10^{-5}	$.65 \times 10^{-5}$
	Total rms	204.	93.5	72.0	$.622 \times 10^{-5}$	2.44×10^{-5}	$.66 \times 10^{-5}$
EXACT MODEL	Tracking Time (days)						
	8	299	93.4	71.9	$.604 \times 10^{-5}$	1.4×10^{-5}	$.62 \times 10^{-5}$
	16	365	94.1	65.8	$.55 \times 10^{-5}$	1.52×10^{-5}	$.42 \times 10^{-5}$
	24	372	96.6	56.0	$.48 \times 10^{-5}$	2.28×10^{-5}	$.23 \times 10^{-5}$
	32	269	104.5	46.7	$.42 \times 10^{-5}$	3.65×10^{-5}	$.13 \times 10^{-5}$
	40	107	114.	42.9	$.402 \times 10^{-5}$	4.26×10^{-5}	$.19 \times 10^{-5}$

Chapter IV will provide strong motivation to try to incorporate them somehow in the estimation scheme. Just how to do this is not clear, however. The main difficulty is that not only are the forces as a function of time unknown, very little information is usually available to even describe the stochastic process from which they are drawn. The main sources of the forces are:

1. The solar pressure force and its interaction with the spacecraft. To a first approximation this can be modeled as a $1/r^2$ force, much like solar gravity itself, except it is non-conservative. Depending on the axial symmetry of the spacecraft, this force will be principally directed outward from the sun. In practice the main source of uncertainty here is in predicting the effective albedo of the spacecraft (particles that are reflected will impart twice the momentum as those absorbed) and its time varying characteristics (as the spacecraft ages, the reflective properties change, usually in a secular manner). See Reference 20.

2. Fluctuations in the solar constant produce stochastic variations about the average solar pressure force described just above. These fluctuations are not very predictable and amount to perhaps .1 to .2%. On a Mariner class spacecraft (845 lbs.) this would produce acceleration variations on the order of 10^{-13} km/sec^2 at Mars distance. See Reference 20.

3. Gas leakage from the jets on the attitude control system have proved the most troublesome in the past. The leakage is due to such phenomena as dust particles trapped in the valves themselves (this changes each time the valves fire which is perhaps once/hr and in general tends to be erratic) and imbalance in the coupling. Each spacecraft flown shows different characteristics in this area. To model it requires careful study of the attitude control sub-system data and is not really tractable, particularly during the real-time mission. See Reference 21.

4. Leakage from the propellant systems and scientific instruments. This item is highly spacecraft dependent and will become more important as spacecraft carry more propellant under high pressure (e.g., for planetary orbiters) and more complex scientific packages (e.g., the infrared spectrometer gas blowdown during Mariner VI and VII encounter blurred the information inherent in the tracking data enough to make it nearly unusable for celestial mechanics purposes). See Reference 22.

5. For solar electric or nuclear propulsion systems, the powered flight phase extends during the entire (or nearly) cruise phase. Fluctuations in the thrust magnitude and direction produce stochastic spacecraft accelerations perhaps as high as 10^{-9} km/sec^2. These are so high as to make the analysis of the situations qualitatively (as well as quantitatively) different from the so-called ballistic cases we are studying and are beyond the scope of this effort. See Jordan (Reference 23)

The preceding list was for background and to convince that although each specific problem had different implications and should be studied with a specific mission and spacecraft in mind, common to all was that, whatever the situation, our current knowledge does not provide an accurate stochastic model for the spacecraft accelerations with which to derive a minimum variance filter. Some structure is needed and a model thought likely for a Viking '75 spacecraft will be assumed:

1. There will be a fairly large unknown but constant force arising from the leakage in the high pressure propellant storage areas. This will be taken as a mean zero, spherically distributed acceleration of standard deviation 10^{-11} km/sec^2 in each of the three orthogonal spacecraft axes.

2. In addition, there will be the random forces from the sources just described. For the baseline model used here, a standard deviation of 10^{-12} km/sec^2 will be assumed and, again, a correlation time of 25 days. Because the random components are not estimated, the magnitude of the acceleration assumed is not important - i.e., an acceleration of twice the standard deviation, for example, will produce twice the effect. The correlation time is important and supporting studies are performed where this is varied.

Results - Case 1 - Medium and Long Arc Doppler Only

The same Viking trajectory as was used for the station location study was selected. Figures 6.6 and 6.7 show behavior over a 30-day period. Displayed is the uncertainty added in the r direction by the presence of the stochastic forces. An independent process is assumed for each of the three orthogonal spacecraft axes. Of course, from the previous analysis, only the component of acceleration in the Earth-probe direction is expected to have much effect and, in the absence of ranging data, only the radial component of the position estimate would be largely affected.

The bottom curve in Figure 6.6 is for a correlation time of 10^4 days, or essentially a constant process. Based on the results of Chapter IV, a contribution to the uncertainty in r of about 50 km is expected (we are using 10^{-12} km/sec^2 instead of 10^{-11} km/sec^2 even for the constant force to facilitate comparison). But just as in the station location case, the analytic computation is not accurate for $\sigma_{\Delta r}$. It rises quickly from the 50 km value, peaking at 200 km after about 10 days of tracking. The analytic formula is a factor of 4 in error.

Even more disturbing is that the errors build rapidly as the correlation time is decreased. The highest contribution observed is for a 1-day correlation time after 3-5 days of tracking - 1500 km. This is a factor of 30 higher than was calculated analytically for the constant acceleration case. Fortunately, as more data is taken, the effect diminishes until after 20 days there is little difference between $\tau = 1$ and $\tau = 25$. In general, $\tau = 5$ appears to be the worst

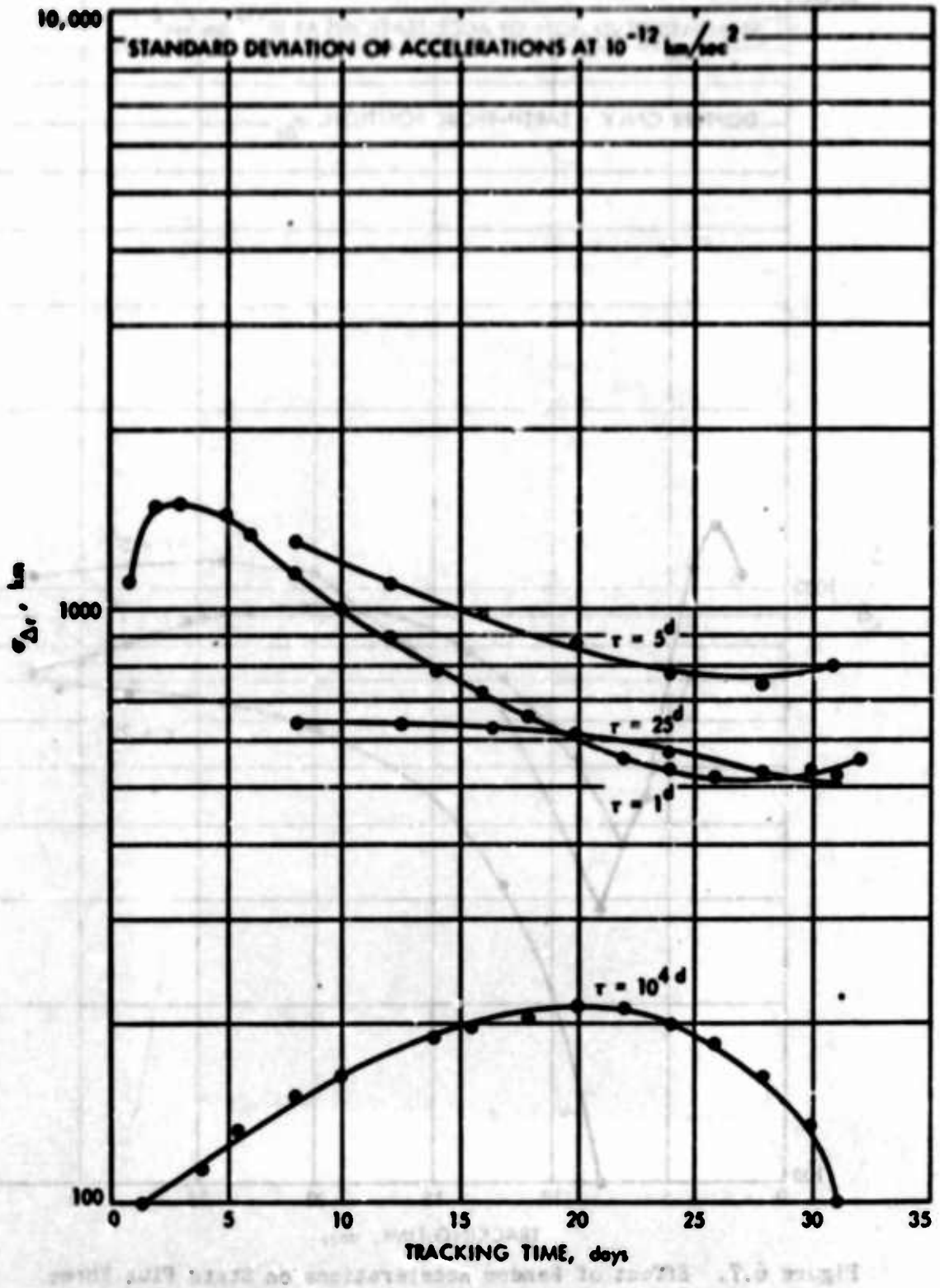


Figure 6.6. Effect of Random Accelerations on State Only Filter versus Tracking Arc, T, For Several Different Correlations Doppler Only; Δr is Earth-Probe Position Component.

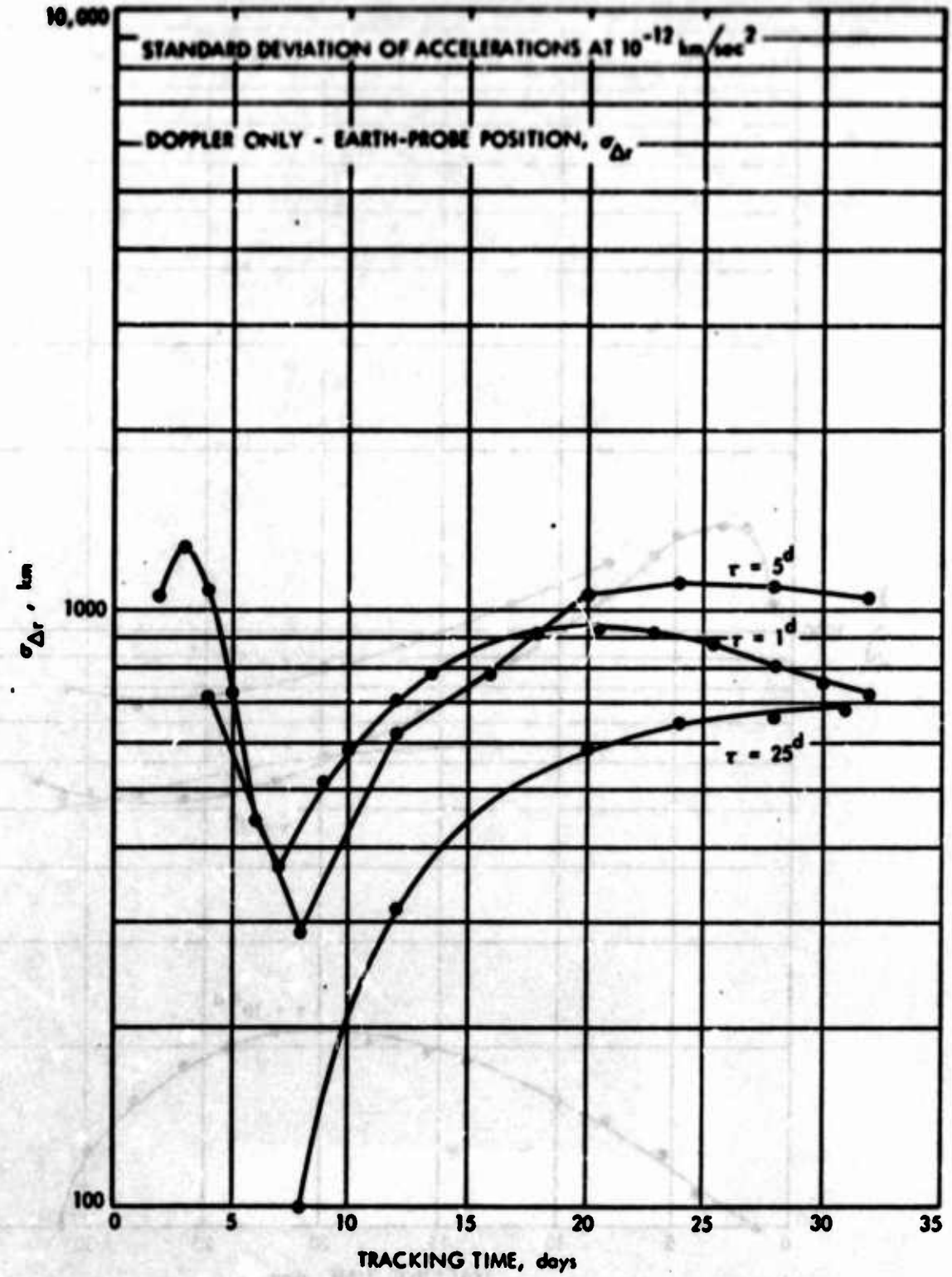


Figure 6.7. Effect of Random Accelerations on State Plus Three Constant Accelerations Filter

correlation time possible.

Figure 6.7 plots the same information for a filter that estimates three constant, orthogonal accelerations assuming an apriori of 10^{-11} km/sec^2. The performance is not much better; in fact, after about 20 days or so, all processes plotted have greater effect on the extended filter than they do on the restricted, state-only filter. This is not to say, however, that the overall performance is not improved since the extended filter handles any constant forces that are present quite well. Where there actually are constant accelerations present at the 10^{-11} km/sec^2 uncertainty level and 10^{-12} km/sec^2 stochastic accelerations, the overall performance (not shown) was better in every case studied.

Figure 6.8 shows the effect of the $\tau = 10^4$ and 5-day accelerations on a state only filter over a much longer six-month tracking interval. The trends are somewhat erratic but the main message is clear-process noise over the data arcs typical for inner planet missions does not produce filter divergence.

Results - Case 2 - Medium Arc with and without Range Data

The preceding showed a remarkable sensitivity to stochastic forces. Although we consider it quite illuminating, the cases studied contain two difficulties:

1. No range data were simulated - it was found in Chapter IV that range data were quite useful in reducing the unmodeled force effects, particularly when the estimate did not have to be propogated very far.

2. To keep the amount of data presented manageable, only the range component of position has been discussed. With ranging data, velocity as well as position errors are induced, but presenting results in all five coordinates involved is

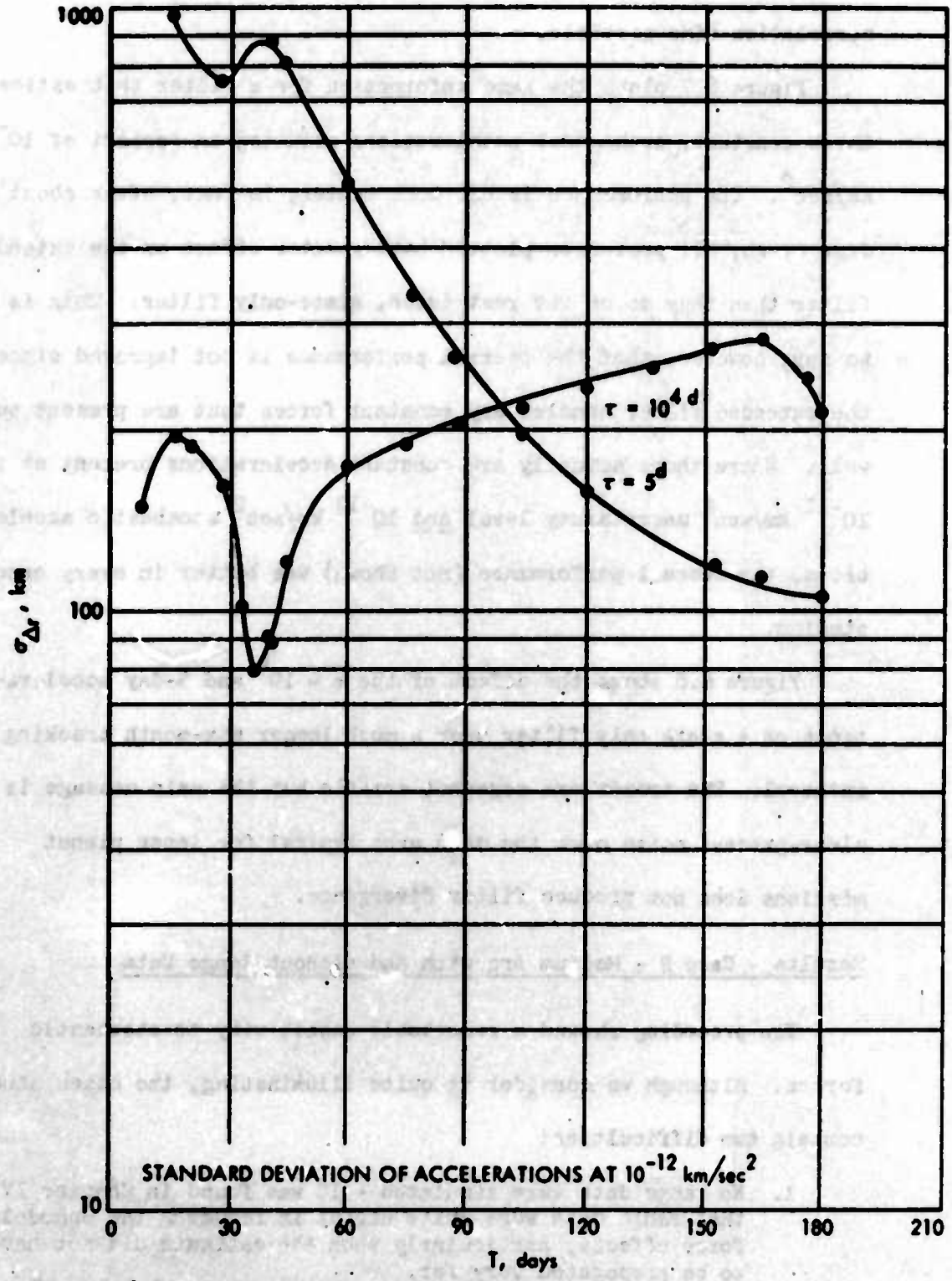


Figure 6.8. Long Term Effect of Random Accelerations on State Only Filter

very awkward.

To circumvent both difficulties, we shall now consider a case including ranging data and display the results in terms of the norm of position of the propogated state error. Specifically, a new Viking '75 trajectory will be considered with epoch and data starting at Mars encounter minus 30 days (E-30d). Tracking data is simulated continuously from three stations until E-10 days at which time the best estimate possible is required in order to calculate a last trajectory correction. Since the mission criteria will require delivery of the spacecraft to a certain Mars relative position, the uncertainties contributed by the random accelerations are propogated to encounter and the square-root of the trace of the position covariance is calculated at that time. In this manner a realistic scalar measure to the 6 dimensional filter performance can conveniently be assigned.

This performance measure is shown plotted in Figure 6.9 for four separate cases which include the combinations of ranging data and not, and extended three constant force solution and not. Each case assumes a correlation time on the stochastic processes of 10 days. Note that the performance measure is not the overall filter performance, it is the effect of the stochastic accelerations only. The following observations may be made:

1. Using doppler data only, the contribution from the forces assumed is roughly 1000 km whether an extended solution is made or not.

2. The addition of ranging data to the state only case improves the performance for very short arcs, but by the time 20 days of data have been accumulated, the error again has risen to approximately 1000 km.

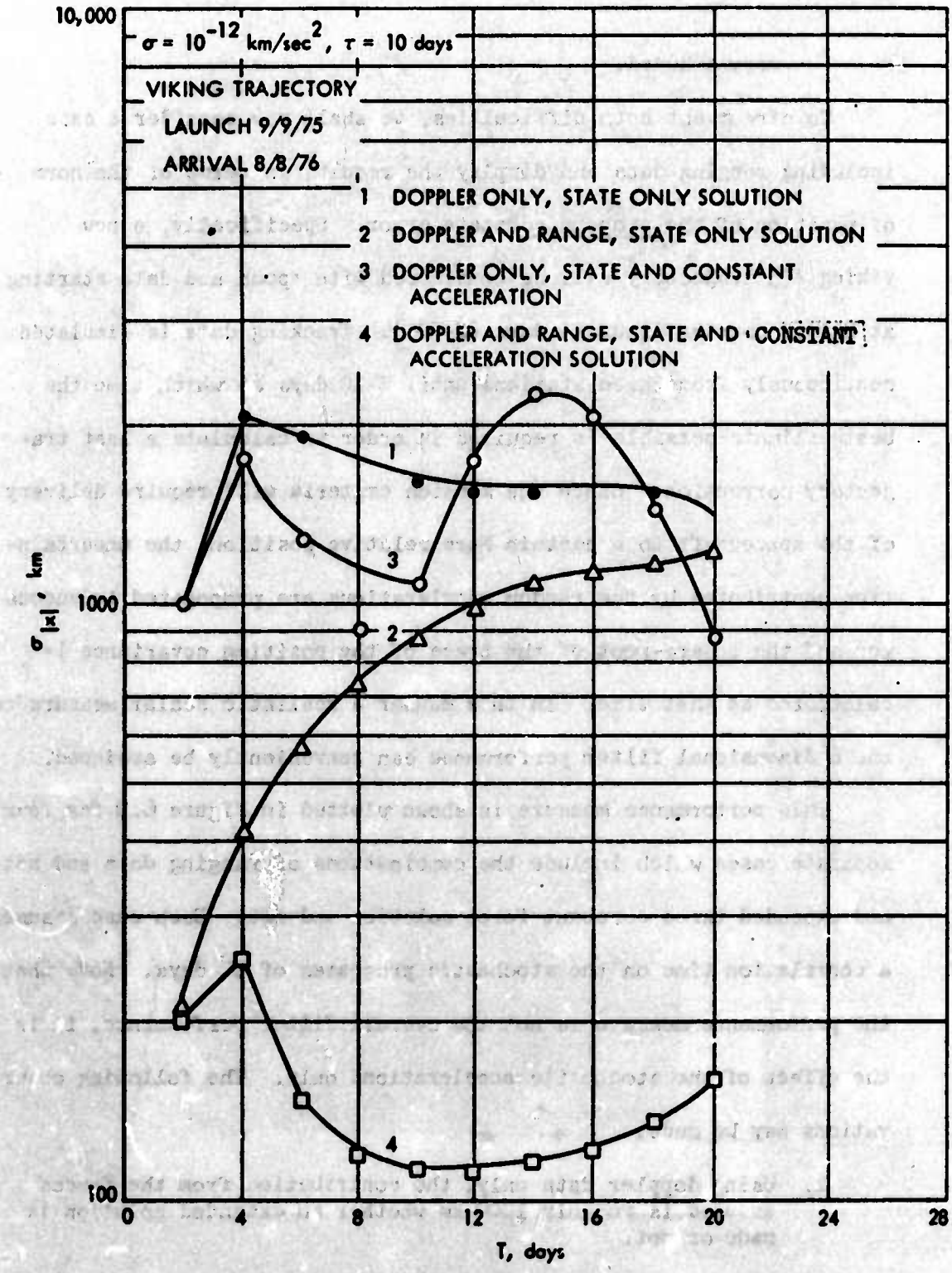


Figure 6.9. Norm of RMS Position Error at Encounter Resulting from Stochastic Accelerations.

3. Using both ranging data and extending the solution produces a markedly superior result. After 20 days of tracking, the error induced is only 170 km.

It is well to check the performance of this later case to make sure that the results are not highly sensitive to the assumed correlation time of 10 days. Figure 6.10 shows results for

$$\tau = \begin{matrix} 1 \\ 5 \\ 10 \\ 25 \end{matrix} \text{ days}$$

using ranging data and the extended solution. As might be expected, performance improves for the longer correlation times since the process is more nearly a constant and hence can be removed efficiently by the extended parameters in the solution. The exception to this is the $\tau = 1$ day case which after 20 days, produces less effect than the $\tau = 5$ day case. This is to be expected because in the limit of zero correlation time, the total power of the stochastic process will be zero. Thus the situation arises where there should be a "worst" correlation time and the results indicate we have found it - at $\tau \approx 5$ days.

Finally, Figure 6.11 presents data yielding the total filter performance including the effects of data noise and the constant forces. There are three curves that can meaningfully be calculated:

① Effects of data noise alone, i.e., the filter solves for three constant forces but there are actually none present.

② Data noise plus constant accelerations, i.e., the filter is the same but this time the constant accelerations actually exist and at the level assumed by the filter ($\sigma = 10^{-11}$ km/sec^2)

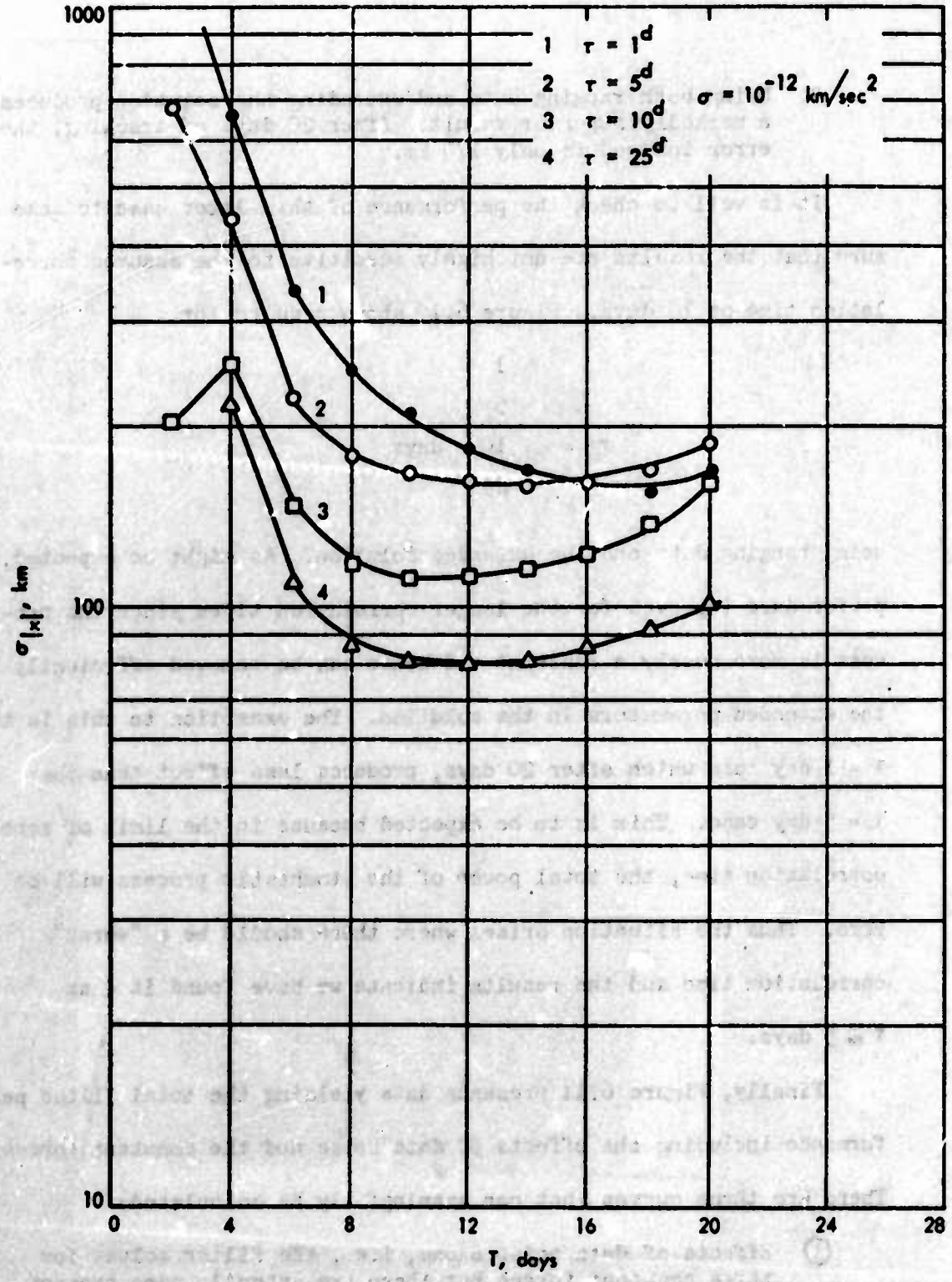


Figure 6.10. Norm of RMS Position Error at Final Time, $\sigma_{|x|}$, Resulting from Stochastic Accelerations.

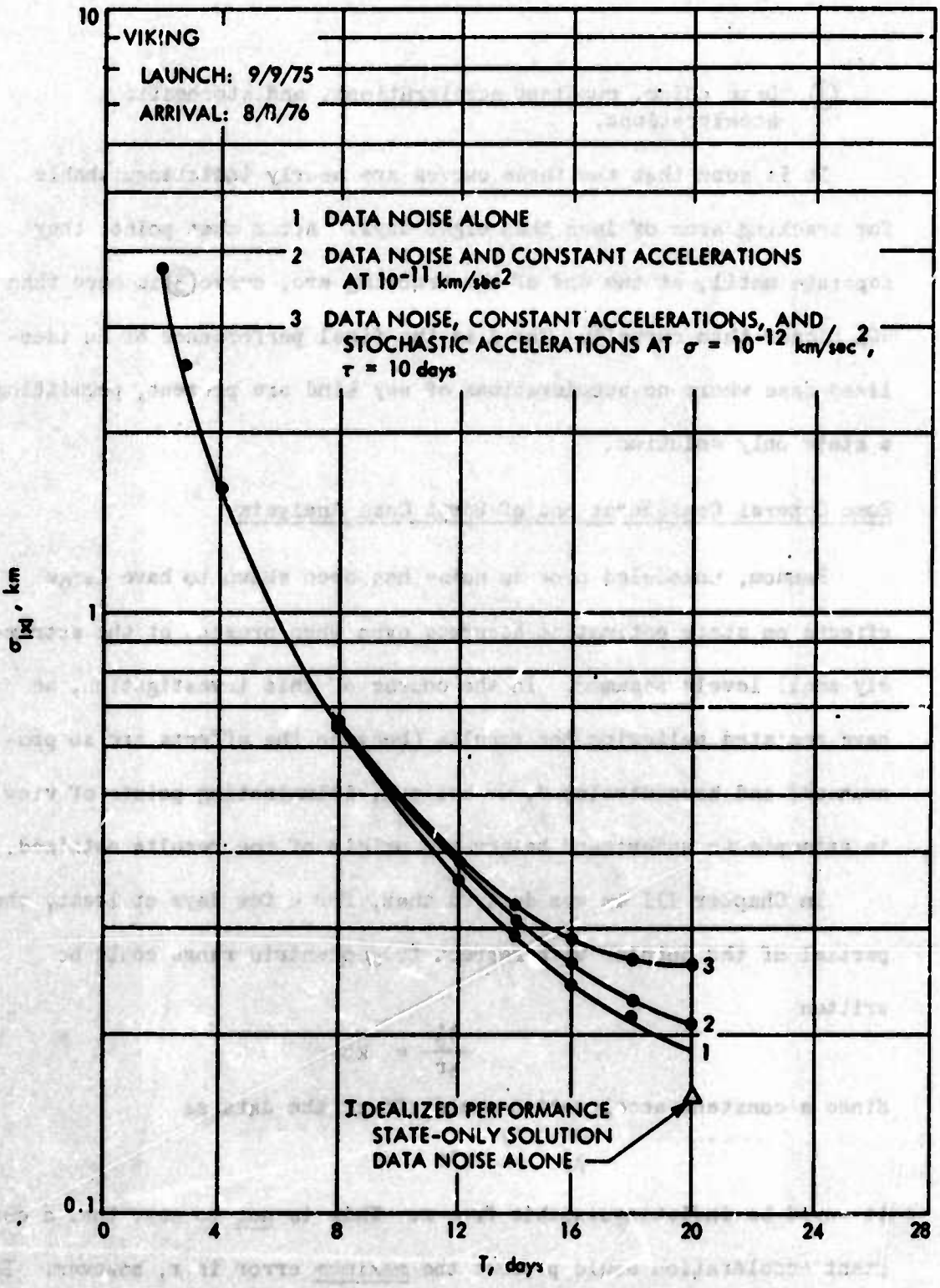


Figure 6.11. Filter Performance, $\sigma_{|x|}$ versus Tracking Arc, T.

③ Data noise, constant accelerations, and stochastic accelerations.

It is seen that the three curves are nearly indistinguishable for tracking arcs of less than eight days. After that point, they separate until, at the end of the tracking arc, curve ③ is more than 40% higher than curve ①. The Δ is the final performance of an idealized case where no accelerations of any kind are present, permitting a state only solution.

Some General Considerations of Worst Case Analysis

Random, unmodeled process noise has been shown to have large effects on state estimation accuracy even when present at the extremely small levels assumed. In the course of this investigation, we have resisted believing the results (because the effects are so pronounced) and have developed, we believe, illuminating points of view in attempts to understand better the origin of the results obtained.

In Chapter III it was derived that, for a few days at least, the partial of the doppler with respect to geocentric range could be written

$$\frac{\partial \dot{\rho}}{\partial r} = kt$$

Since a constant acceleration would affect the data as

$$\Delta\dot{\rho} = k't$$

it would be indistinguishable from r. This is not to say, that a constant acceleration would produce the maximum error in r, however. In order to see this, suppose that a variable, x_1, is being estimated whose data partial over the available data set is orthogonal to every other data partial of the remaining variables being estimated. This

partial is shown in a two-dimensional sub-space of the m dimensional space spanned by the m data points being taken (Figure 6.12). In the Householder notation used, R will have the form

$$\begin{bmatrix} r_{11} & 0 & 0 & \cdots & 0 \\ & r_{22} & r_{23} & \cdots & r_{2n} \\ & & \ddots & & \vdots \\ & 0 & & & r_{nn} \\ & & & & \\ & & 0 & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} z_x \\ \cdots \\ z_z \end{bmatrix}$$

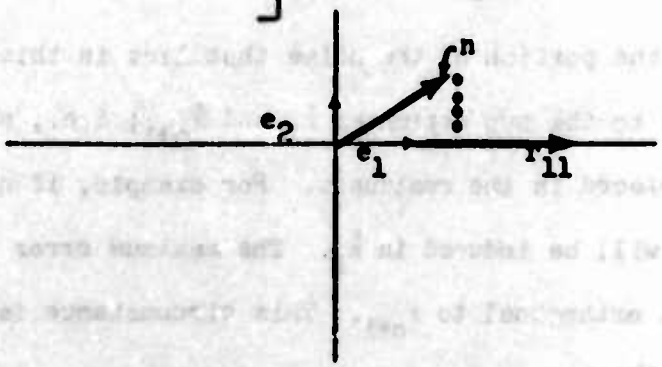


Figure 6.12. Maximum Error Occurs when n lies along r_{11}.

In this circumstance, the error in $\hat{x}_1$ will be simply the dot product of n on the x_1 partial. If, for example, the worst possible noise of a given norm was sought, it would obviously be that n that was co-linear with the original data partial.

Suppose, however, that an additional parameter were added to the estimate list, whose partial was contained within the e_1, e_2 sub-space just described. Call that partial r_{n+1} - it is shown in Figure 6.13.

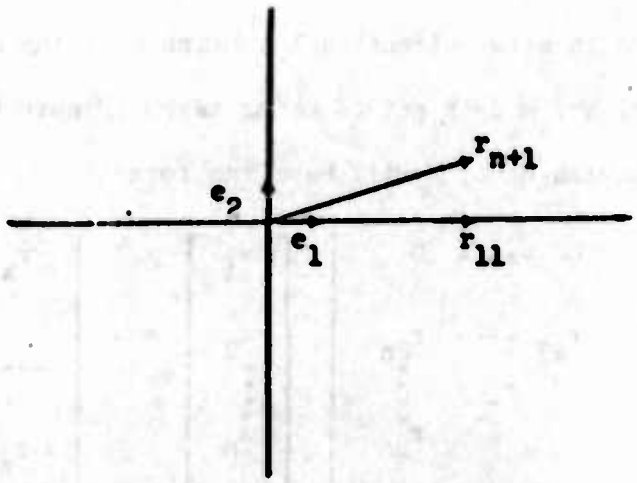


Figure 6.13. When the data partials are not orthogonal, the maximum error does not occur when n lies along n_1.

Now the portion of the noise that lies in this sub-space must be allocated to the two estimates $\hat{x}_1$ and $\hat{x}_{n+1}$; i.e., none of the noise will be placed in the residuals. For example, if n lies along r_{n+1}, no error will be induced in $\hat{x}_1$. The maximum error in $\hat{x}_1$ is induced when n is orthogonal to r_{n+1}. This circumstance is shown in Figure 6.14.

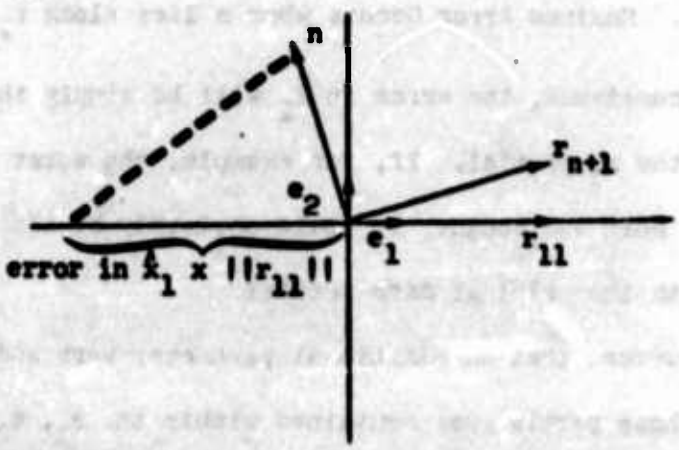


Figure 6.14. Maximum Error in $\hat{x}_1$ occurs when n is orthogonal to r_{n+1}.

The data partials being used to determine the 6 state components are highly correlated. Going back to the analysis of Chapter III,

compare

$$\tau$$

with $\sin \tau$

and $\tau \cos \tau$

These functions, and hence the data partials are very similar - hence the maximum error will not be caused by a noise vector in the direction of the original partial.

Consider a second point of view. Suppose a standard least squares estimator is employed:

$$\hat{x} = (A^T A)^{-1} A^T Z$$

and suppose that the A matrix is given by

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 5 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{\partial z}{\partial x_1} & \frac{\partial z}{\partial x_2} \end{bmatrix}$$

Now if

$$z = Ax + n$$

and defining

$$F \triangleq (A^T A)^{-1} A^T$$

then

$$\hat{x} - x = Fn$$

The F matrix is readily seen to be

$$F = \frac{1}{12} \begin{bmatrix} 5 & 5 & -3 & 5 \\ -1 & -1 & 3 & -1 \end{bmatrix}$$

Suppose the error in $\hat{x}_1$, only is of direct interest. It is desired

to calculate the largest error in this variable if, say, each element of n were bounded between ± 1, but was arbitrary otherwise. If

$$n = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

were selected (i.e., a noise signature that "looked" like the original data partial), the $\hat{x}_1$ error would be unity. If however

$$n = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

were selected, an error of 1.5 would result, the largest possible for the problem posed.

To return to the process noise situation under study, suppose the worst possible unmodeled acceleration signature is desired with the only constraint being that the acceleration be a bounded variable. Consider the row of the F matrix of the coordinate being studied (r would be ideal) as a continuous function. The error induced would be given by

$$\Delta\hat{x} = \int_0^T f(\tau)\, \delta z\, (\tau)\, d\tau$$

where δz is the perturbation in the data. If the approximation that (assuming doppler data only) $\delta z\, (\tau) = \int_0^\tau a\, (u)\, du$ is made where a (u) is the bounded acceleration (this would be an excellent approximation on a deep space trajectory for a month or so). The problem becomes:

Extremalize $\Delta\xi$ subject to

$$| a(\tau) | \leq K$$

This is a straightforward Calculus of Variations problem. Define the Hamiltonian

$$H(\delta z, a, \lambda, t) = f(t)\,\delta z(t) + \lambda(t)a(t)$$

Then minimizing with respect to a (t), we find

$$a(t) = -K \operatorname{sgn} \lambda(t)$$

Since on an extremal arc

$$\frac{d\lambda}{dt} = -\frac{\partial H}{\partial z} = -f(t)$$

Then

$$\lambda(t) = -\int_0^t f(\tau)\,d\tau + C$$

Because there are no end-point constraints

$$\lambda(\tau) = 0$$

and

$$\lambda(t) = \int_t^\tau f(\tau)\,d\tau$$

thus

$$a(t) = -K \operatorname{sgn} \int_t^\tau f(\tau)\,d\tau \quad * \qquad (1)$$

This is the answer in the continuous formulation - the discrete version yields highly similar answers. Figure 6.15 plots the discrete

---

*We are following a straightforward application of Pontryagin's Maximum principle as outlined in D. M. Wiberg's class notes for Engr. 222-C, Winter Quarter, 1969. The development is part of a separate project being carried out by James McDonell and this author. It is beyond the scope of this thesis, but we wish to apply some of the results here and will describe the procedures being used briefly.

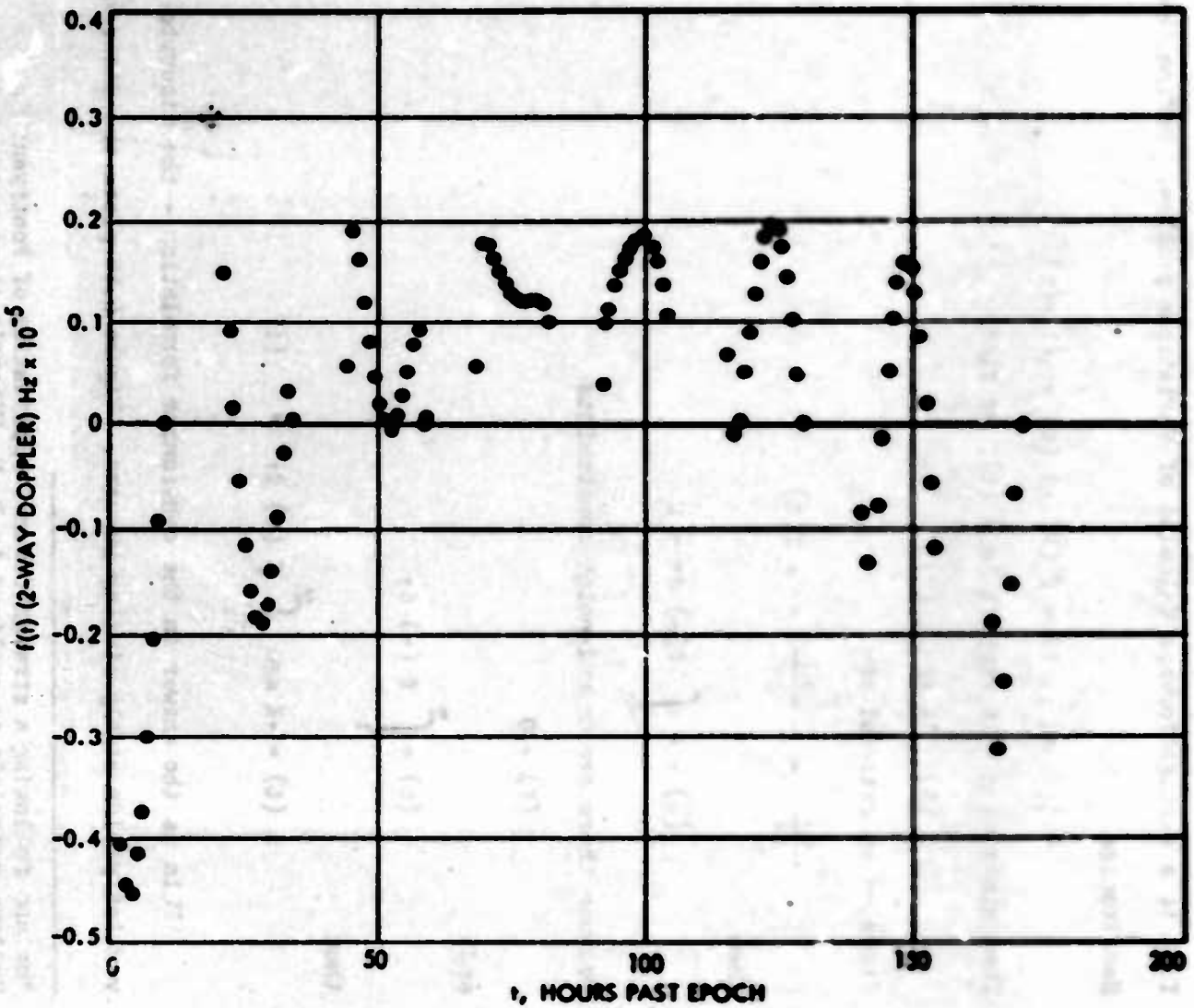


Figure 6.15. Discrete F Function with respect to r coordinate for eight days Doppler Tracking.

f functions for the r estimate for the first eight days of doppler data on the Viking trajectory of Case I just above. Note how dissimilar it is from the simple k't structure of the original partial. If equation (i) is executed by summing the points of this plot starting at the right-hand side, there are two reversals of sign as the running sum is accumulated. Thus, the worst case acceleration is one that starts at +K, reverses once to -K, and finally reverts to +K again as T interval is traversed. This is hardly a constant acceleration. This example lends credence to the covariance results which determined that a 5-day correlation time is the worst case and gives some insight as to the origin of this result.

CHAPTER VII.

FURTHER CONSIDERATIONS IN DETECTING UNMODELED PARAMETERS

Much of the preceding material has dealt with filter response to unmodeled or non-precisely modeled parameters and processes. The presence of un-modeled parameters can seriously degrade state estimates as has been shown. Solving for them, even non-optimally, can often be very helpful as was detailed in Chapter VI. There is a price for this solution, though, and that is that the state estimates are more sensitive to data noise. That is, as more parameters are included in the solution, the filter becomes more broadly tuned, and its discrimination powers between signal and noise are correspondingly less. Thus, if a parameter is thought not to be present (at a significant level) but it is feared that it might be, the "fear" can be covered (often) with a solution but given that the "thought"was correct, the covariance of state errors will be higher if a solution is made than if it is not.

These remarks can be put on a quantitative basis by employing the generalized consider option of Chapter V. For example, filter performance can be calculated for the situations where constant non-gravitational forces are either solved-for or not and are present or not. Table 7.1 displays the results for 4 separate cases for two different spans of doppler tracking data:

1. Filter estimates state only; forces not present.
2. Filter estimates state and three forces; but forces are not

Case No.	Solve For Forces	Forces Present	Uncertainty in Radial Direction (km)	
			8 Days of Tracking	32 Days of Tracking
1	No	No	415	70
2	Yes	No	760	84
3	Yes	Yes	1024	511
4	No	Yes	1515	1110

A priori force uncertainty of 10^{-11} km/sec^2 was used where appropriate in both filter and actual situation.

Table 7.1. RMS Uncertainty in Range Direction for Various Combinations of Physical Model and Filter Design.

present.

3. Filter estimates state and three forces; forces do exist.

4. Filter estimates state only; forces exist.

The variable recorded is the r direction, the main state parameter affected by unmodeled forces. In general, while the true filter performance follows the numbering system just adopted, the specific point being made is that Case 2 shows degraded performance from Case 1.

The two parameters classes dealt with in previous chapters (station locations and non-gravitational forces) present the very difficulties for real time orbit determination of which we are speaking. Spacecraft are designed carefully to minimize translational forces being spuriously generated by the attitude control, propulsion, and other systems for this very reason. They can develop (and have, see Ref.26) leaks at any time, however. As for station locations, these parameters are carefully determined using data from previous missions-there is, however, the not very remote possibility that the solutions are not as accurate as is thought, the propogation model is inaccurate, or that the universal time calibrations are inaccurate for the new time spans being used.

These factors combine to motivate the search for computational machinery to augment existing orbit determination methods that aid the analyst in performing formal or informal hypothesis testing with regard to his solutions. The data which he has at his disposal are 1) the temporal or data set dependent behavior of the solutions themselves and 2) the data residuals after the solution(s) has been made. In this chapter we wish to suggest a simple algorithm to serve as a

"detectability index", a means of deciding to what extent an unmodeled parameter, if present, would be noticed by inspecting the data residuals, or more simply, the weighted sum-of-squares of the residuals.

A Considered Residuals Option

Calculating the signature impressed on the data by a hypothetical parameter is not directly useful in pursuing the goal of detecting unmodeled parameters. To look for this signature, one would have to inspect the data before the fit, or solution, were made and this data would contain the signatures from all other state parameters as well. We must be content looking for this desired signature after it has passed through the filter to be used. Of course its character will be changed by this operation - formally, what will remain is the portion orthogonal to the column space of the partial matrix used to design the filter.

For example, consider a data set linearly related to the estimated state, x, and unmodeled parameters, y, as follows

$$z = Hx + By + \epsilon \tag{7.1}$$

Clearly, y adds By to z. If the least squares filter,

$$\hat{x} = (H^T \Lambda_\epsilon^{-1} H)^{-1} H^T \Lambda_\epsilon^{-1} z \triangleq F z$$

is employed and $H\hat{x}$ is subtracted from z to form the residual vector, then

$$\delta z \triangleq z - H\hat{x} = (I - H F) By + (I - H F) \epsilon \tag{7.2}$$

The first half of the right hand portion of 7.2 are the residuals due to the presence of y.

In any given problem, having on hand this filtered signature to compare with the residual patterns actually obtained might be quite

useful. It has the disadvantage of being unwieldy, unfortunately, since its application would require plotting of these signatures for each parameter being "feared" and for each data set (or span) used for the various solutions made. The curse of too much information lurks nearby this suggestion. Something of more facile application would be to calculate the increase in the weighted sum-of-squares due to the existence of an unmodeled parameter. Using 7.2 this is easily seen to be

$$\delta z^T \delta z \text{ due to } y = y^T B^T (I - H F)^T (I - H F) B y$$

$$\text{or} \qquad (7.3)$$

$$= \text{trace } (I - H F) B y y^T B^T (I - H F)^T$$

Unfortunately, either form of (7.3) involves literal manipulations with matrices of very inconvenient dimensions (for example the number of rows of B is equal to the dimension of the data vector). The Householder transformation technique, which automatically selects a new basis for the partial derivative matrix can be conveniently employed to solve the dimensionality problem imbedded in 7.3.

We begin the formal development by once again employing 7.1 recast as a square-root filtering problem as described in the Appendix and extended in Chapter V.

$$\begin{bmatrix} H & \vdots & B \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} + \begin{bmatrix} n_{\hat{x}} \\ n_{\hat{y}} \\ n_z \end{bmatrix} = z = \begin{bmatrix} H & \vdots & B \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \epsilon \qquad (7.4)$$

As before, a solution for 7.4 is desired that minimizes the norm of the weighted residual vector. The y variables are again the unmodeled parameters which may or may not be present - their estimates will be

left at zero in either case. Pre-multiplying 7.4 by the square-root weighting matrix and then performing the Householder transformation, recall the results of Chapter V and re-write as:

$$\begin{bmatrix} \hat{x}-x \\ \hat{h}_y \\ \hat{h}_z \end{bmatrix} = \begin{bmatrix} R_x^{-1} & R_{xy}y \\ R_y y & \\ 0 & \end{bmatrix} + \begin{bmatrix} R_x^{-1} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} P W^{\frac{1}{2}} \epsilon \qquad (7.5)$$

Note that the presence of y results in a contribution to the residuals only in p of the elements in terms of the new basis which has been created (where p is the dimension of y). It is convenient to think of y as a zero mean random variable uncorrelated with x (as was done in the development of the consider option) and compute the covariance of $\hat{x}-x$, $\hat{h}_y$, $\hat{h}_z$. Post-multiplying 7.5 by its transpose and taking ensemble averages.

$$\begin{bmatrix} \Lambda_{\hat{x}} & \Lambda_{\hat{x}\hat{h}_y} & \Lambda_{\hat{x}\hat{n}_y} \\ S_{ym} & \Lambda_{\hat{n}y} & \Lambda_{\hat{n}y\,\hat{n}z} \\ & & \Lambda_{\hat{n}z} \end{bmatrix} = \begin{bmatrix} R_x^{-1} & R_{xy} \\ & R_y \\ & 0 \end{bmatrix} \Lambda_y \begin{bmatrix} R_{xy}^T R_x^{-1T}, & R_y^T, & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} R_x^{-1}R_x^{-1T} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \qquad (7.6)$$

The upper left partition of 7.6 is the covariance of state estimate errors and is identical to 5.3. Writing out the lower right 2 x 2 partition, the covariance of residuals:

$$\begin{bmatrix} \Lambda_n^{\Delta} \\ \\ \end{bmatrix} = \begin{bmatrix} R_y \Lambda_y R_y^T + I & 0 \\ 0 & I \end{bmatrix} \begin{matrix} \} \ p \\ \} \ m-(p+n) \end{matrix} \qquad (7.7)$$

where p is the dimension of the consider parameter vector

m is the dimension of the data vector

n is the dimension of the consider solution vector

and since

$$x^T x = \text{trace } x\, x^T$$

it is concluded that the trace of (7.7) is the expected value of the weighted sum-of-squares of the residual vector. Note that if Λ_y is zero, this is simply m-n, the classical least-squares result for uniformly weighted data. It is interesting that this is also the result for an arbitrary weighting of the data, W, as was employed here. More compactly:

$$E\left[\text{weighted sum-of-squares}\right] \triangleq \text{sos} = (m-n) + \text{tr } R_y \Lambda_y R_y^T \qquad (7.8)$$

Equation (7.8) can be computed efficiently and is a useful result in itself but still does not get at the notion of detectability. It would be more desireable to compute, say, the per-cent increase in sos, per unit change in the estimate due to an unmodeled parameter. Equation (7.8) is cast statistically but it is obviously possible to write the increase in the sos in terms of a particular y, i.e.,

$$\Delta\text{sos} = \text{tr } R_y\, y\, y^T R_y^T$$

For example, if p were 2 then

$$\Delta\text{sos} = r_{11}^2\, y_1^2 + (r_{12}^2 + r_{22}^2)\, y_2^2 + 2\, r_{11}\, r_{12}\, y_1 y_2 \qquad (7.9)$$

where r_{ij} is the i, j element of R_y, an upper triangular matrix.

Generalizing (7.9) it is easily shown that the increase in the sos due to the presence of a single parameter y_i is

$$\Delta sos_i = y_i^2 \sum_{j=1}^{i} r_{ji}^2 \qquad (7.10)$$

The cross-product terms do not appear in (7.10) since each parameter is treated separately - that is, in postulating that y_i has a non-zero magnitude, it is implicitly assumed that the rest of the parameters are zero.

As for the sensitivity of the estimate to the unmodeled parameter, this is readily seen to be (from 7.5)

$$\Delta\hat{x} = R_x^{-1} R_{xy} y$$

Again considering only one element of y, y_i,

$$\Delta\hat{x} = R_x^{-1} r_{xy}(i)\, y_i \qquad (7.11)$$

where $r_{xy}(i)$ is the ith column of R_{xy}.

Finally, by combining (7.10) and (7.11), the following suggested inverse detectability index is established.

$$d_i \triangleq R_x^{-1} r_{xy}(i) / \sqrt{\sum_{j=1}^{i} r_{ji}^2/(m-n)} \qquad (7.12)$$

In words, (7.12) gives the vector change in the estimated parameters divided by the per cent increase in the rss weighted residuals (over that expected statistically from noise alone) due to the ith unmodeled parameter. Thus, speaking qualitatively, in a real fitting problem, even though it can be demonstrated that the estimate is sensitive to y_i, if d_i is small, and the residuals conform well to that expected from noise alone, it can be concluded at a glance that

y_i is not present at a damaging level.

This index should work well in validating apparently good solutions in a real-time situation. If the residuals do not show structure, the sos is nearly what is expected, and all the d_i's are adequately low, then it could be concluded with reasonable assurance that nothing untoward has occurred and the modeling has been faithfully executed. In pre-analysis, the index might be useful in establishing what data sets are appropriate for the execution of the validation process. In general, the d_i's should become lower with increased amounts of data.

If, on the other hand, the data does not fit well and there are several parameters that could be causing the difficulty, it would be useful to have a convenient means of investigating which of these parameters (if any) when included in the solution would decrease the sos by significant amounts. It would be desireable to have an algorithm to do this without explicitly performing the various solutions*. To find the desired algorithm, first break the consider parameter vector, y, into three sub-classes:

$$y \triangleq \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

In terms of these newly defined y_i's, we wish to compute the reduction on the sos that would be obtained if y_2 were added to the estimation

*That this could be done was first pointed out to the author by S. R. McReynolds in a private communication.

list, but the estimates of y_1 and y_3 were still left at zero. Again, y_2 is restricted to having unity dimension and only the possibility of adding a single new variable to the estimation list is investigated. The following equation must be solved:

$$\begin{bmatrix} R_y & R_{xy_1} & r_{xy_2} & R_{xy_3} \\ 0 & R_{y_1} & r_{y_1y_2} & R_{y_1y_3} \\ 0 & 0 & r_{y_2} & R_{y_2y_3} \\ 0 & 0 & 0 & R_{y_3} \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{y}_1=0 \\ \hat{y}_2 \\ \hat{y}_3=0 \end{bmatrix} + \begin{bmatrix} \hat{n}_x \\ \hat{n}_{y_1} \\ \hat{n}_{y_2} \\ \hat{n}_{y_3} \end{bmatrix} = \begin{bmatrix} z_x \\ z_{y_1} \\ z_{y_2} \\ z_{y_3} \end{bmatrix} \tag{7.13}$$

for $\hat{x}$, $\hat{y}_2$ such that the norm of $\hat{n}$ is minimized.

Writing (7.13) as a set of equations:

$$\begin{aligned} R_x \hat{x} + r_{xy_2} \hat{y}_2 + \hat{n}_x &= z_x \\ r_{y_1y_2} \hat{y}_2 + \hat{n}_{y_1} &= z_{y_1} \\ r_{y_2} \hat{y}_2 + \hat{n}_{y_2} &= z_{y_2} \\ \hat{n}_{y_3} &= z_{y_3} \end{aligned} \tag{7.14}$$

From the structure of 7.14:

1. There is no freedom to alter $\hat{n}_{y_3}$ from that specified.
2. Whatever the choice of $\hat{y}_2$, $\hat{x}$ can be adjusted so that $\hat{n}_x$ may be zero.
3. Therefore the problem resolves to one of choosing $\hat{y}_2$ so

that the $\left\| \begin{matrix} \hat{n}_{y_1} \\ \hat{n}_{y_2} \end{matrix} \right\|$ is a minimum.

From (7.14), it is evident that

$$\hat{n}_{y_1}^T \hat{n}_{y_1} + \hat{n}_{y_2}^2 = (z_{y_1}^T z_{y_1} + z_{y_2}^2) - 2(z_{y_1}^T r_{y_1 y_2} + z_{y_1} r_{y_2})\hat{y}_2 + (r_{y_1 y_2}^T r_{y_1 y_2} + r_{y_2}^2)\hat{y}_2^2 \qquad (7.15)$$

Minimizing (7.15) with respect to $\hat{y}_2$ yields

$$\hat{y}_2 = \frac{(z_{y_1}^T r_{y_1 y_2} + z_{y_1} r_{y_2})}{(r_{y_1 y_2}^T r_{y_1 y_2} + r_{y_2}^2)}$$

and the <u>decrease</u> in the sos is simply

$$\Delta sos = \frac{(z_{y_1}^T r_{y_1 y_2} + z_{y_1} r_{y_2})^2}{(r_{y_1 y_2}^T r_{y_1 y_2} + r_{y_2}^2)} \qquad (7.16)$$

As an application example, suppose that several stations are taking data and through a blunder a single component of one of the station locations is seriously in error and none of these locations are included in the solution. All of the data residuals will be affected by the error as a best compromise is found by the lease-square algorithm. The actual error component could be isolated very quickly, however, by executing (7.16) for each consider parameter, since only that parameter would be completely successful in removing the structured residual patterns that have been produced.

A Numerical Example

An 1108 Computer program was mechanized to calculate the sos changes and the detectability index for several possible modeling errors. Again as in the preceding chapter, the program was designed to use the partials generated by the DPODP. To continue the theme of the earlier chapters, non-gravitational forces and the station location errors are selected for the numerical examples.

Figure 7.1 plots the increase in the sos over that expected, from data noise alone, i.e.,

$$\frac{\Delta sos}{E\left[sos\right]} = y_i^2 \sum_{j=1}^{i} r^2_{ji}/(m-n)$$

assuming a solution for the six components of the trajectory was made.

The DPODP models the non-gravitational forces as components, not in the Earth-relative plane-of-the-sky, but along the spacecraft axes. Since the roll axis is oriented along the sun-probe line, this becomes one component, a_r. The other two componets are a_y which is in the ecliptic but perpendicular to the direction of a_r and a third direction which is perpendicular to the ecliptic, a_x. The results could be related better to the preceding chapters if the plane-of-the-sky coordinate system were used, but we are interested in detecting general classes of parameters here-not just those which are well understood. This rather arbitrary coordinate system will suffice. The station location errors are denoted in the coordinate system being used, r_s and λ.

From Figure 7.1, as expected the per cent change in the sos rises as more tracking data are included. The formal principle at work here

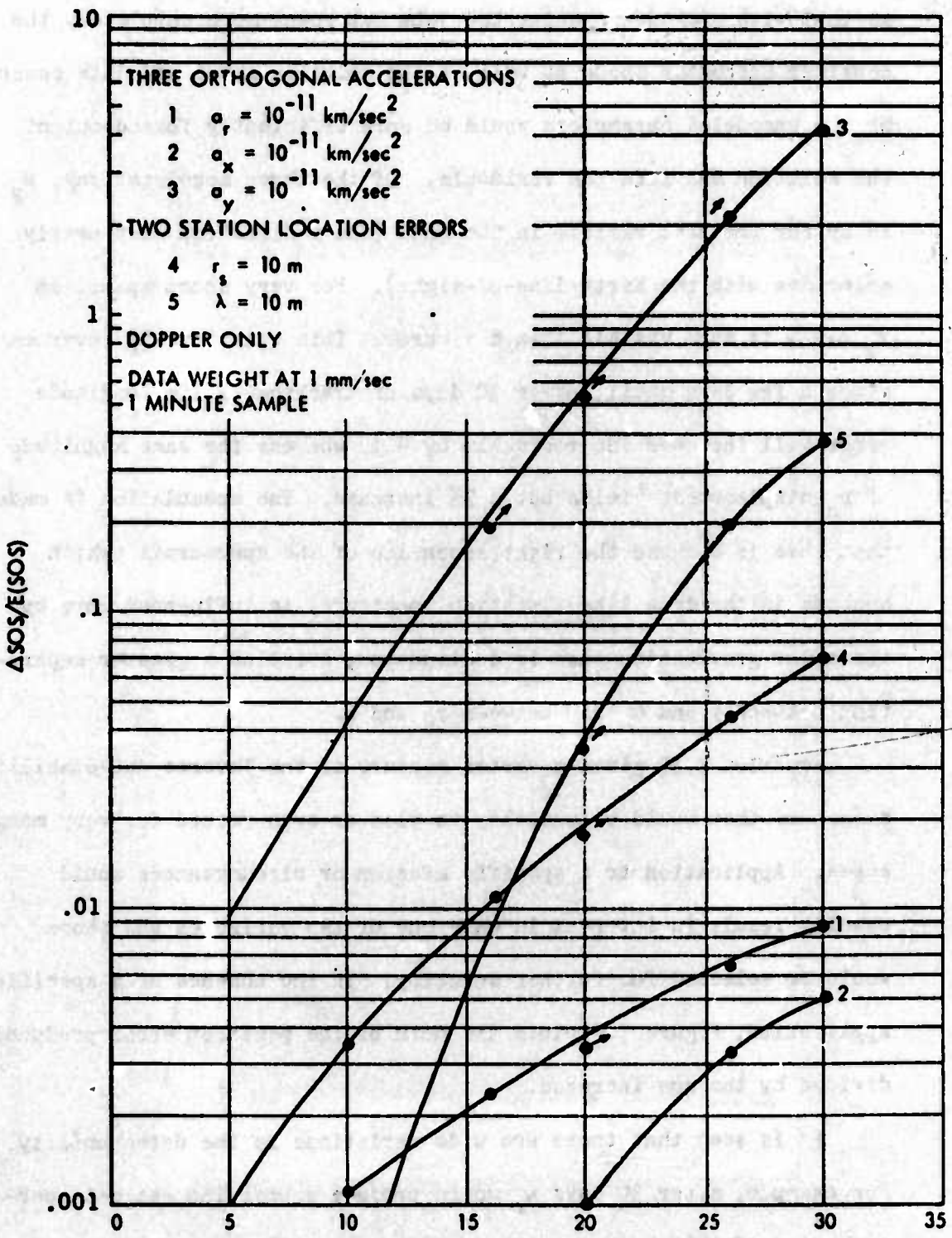


Figure 7.1. Change in sos divided by that expected versus Tracking Arc.

is that with additional data, the data set spans more completely the consider parameter space as well as the solution space and data caused by the unmodeled parameters would be more efficiently forced out of the solution and into the residuals. Of the three accelerations, a_y is by far the most visible in the data (the y direction most nearly coincides with the Earth-line-of-sight). For very short spans, an r_s error is more visible than a λ error. This relationship reverses after a few days until, after 30 days of tracking, a 10m longitude error will increase the residuals by 40%, whereas the same magnitude of r_s displacement yields but a 7% increase. The speculation is made that this is because the right ascension of the spacecraft (which appears in the data like a station longitude) is influenced more by the solar gravitation than is declination, creating a greater separation between λ and α than between r_s and δ.

Equation 7.12 gives a vector measure of the inverse detectability, a measure that would be unwieldy to plot or even record for very many cases. Application to a specific mission or circumstances would usually result in interest in only one or two variables and those would be selected for further scrutiny. In the absence of a specific application, Figure 7.2 plots the norm of the position error produced divided by the sos increase.

It is seen that there are wide variations in the detectability. For example, after 30 days a_r would produce a position estimate perturbation of 3600 km before the sos rose by 100%. For the same sos change, a_x would produce only a 180 km error. Again the relative relationship between r_s and λ effects changes as more data is added.

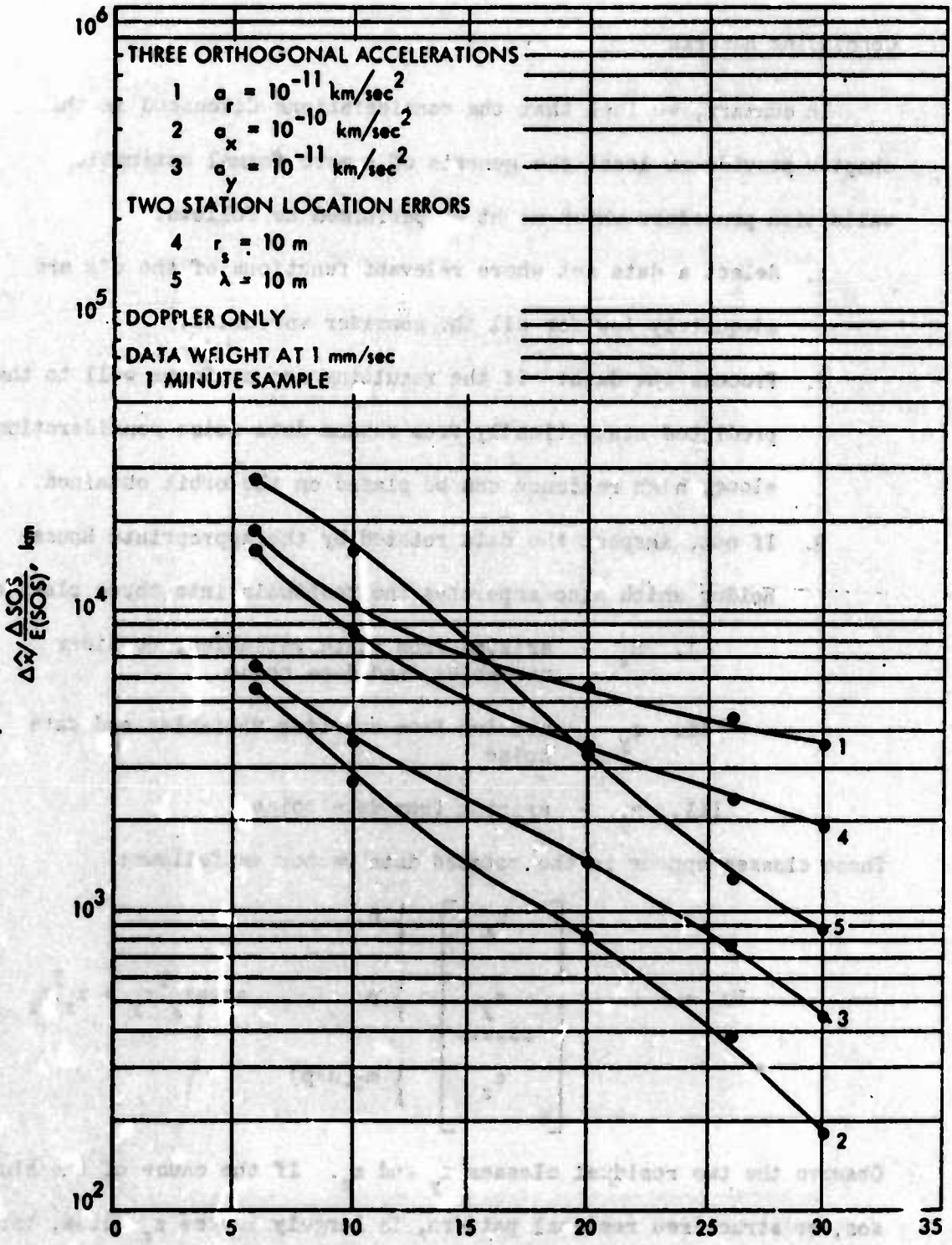


Figure 7.2. Injection Position Error Per Unit sos Increase versus Tracking Time.

Concluding Remarks

In summary, we feel that the considerations discussed in this chapter provide at least the genesis of a more formal estimation validation procedure which might be performed as follows:

1. Select a data set where relevant functions of the d's are adequately low for all the consider variables.
2. Process the data: if the resulting sos conforms well to that predicted statistically from random data noise considerations alone, high reliance can be placed on the orbit obtained.
3. If not, inspect the data rotated by the appropriate Householder which also separates the residuals into three classes:

   I. z_x - arising from state variables, consider variables, and data noise

   II. z_y - arising from consider variables and data noise

   III. z_y - arising from data noise.

These classes appear in the rotated data vector as follows:

$$Pz = \begin{bmatrix} z_x \\ \hline z_y \\ \hline z_z \end{bmatrix} \begin{matrix} \} n \\ \} p \\ \} m-(n+p) \end{matrix} \qquad sos = z_y^T z_y + z_z^T z_z$$

Observe the two residual classes z_y and z_z. If the cause of the high sos, or structured residual pattern, is largely in the z_z class, this means that what is causing the difficulty has not been modeled even in the consider parameters, let alone the solution parameters. In this

case there is little that can be done other than reaching the unsatisfying conclusion that the difficulty is at least in part arising from considerations not even thought of.

4. If the high residuals are largely contained in the z_y class, the consider parameters span the difficulty. Execute the test algorithm (Equation 7.16) to determine if one or perhaps a few consider parameters are present at a high value. Once the culprit(s) has been isolated, add it to the estimation list and re-perform the procedure.

APPENDIX

I. Orbit Determination

Orbit determination is the process of using imperfect observations which are functionally dependent on the state of the observed body to form an estimate of the body's trajectory which is best in some defined sense. We shall be speaking of trajectories subject primarily to gravitational force fields and in particular, spacecraft trajectories. Further, we shall be concerned primarily with earth-based observations of the spacecraft; the most useful of these are tracking-station to spacecraft range and range rate. They will be discussed in more detail in the following sub-section.

The basic process of using the measurements to form the estimate is shown schematically in Figure A.1. Here, our point of view is to obtain an estimate of the initial conditions, or state, and then through integration of these estimates form the complete estimated trajectory. It is often convenient, though by no means a requirement, to visualize the initial conditions as the position and velocity of the spacecraft at that epoch. More generally, the initial conditions will include both the spacecraft state and other parameters that affect either the propogation of spacecraft state and/or the measurements themselves. To be more explicit we define:

$$q \triangleq \begin{array}{l}\text{total parameter}\\ \text{vector being}\\ \text{estimated}\end{array} \triangleq \begin{bmatrix} x(0) \\ v \\ s \end{bmatrix}$$

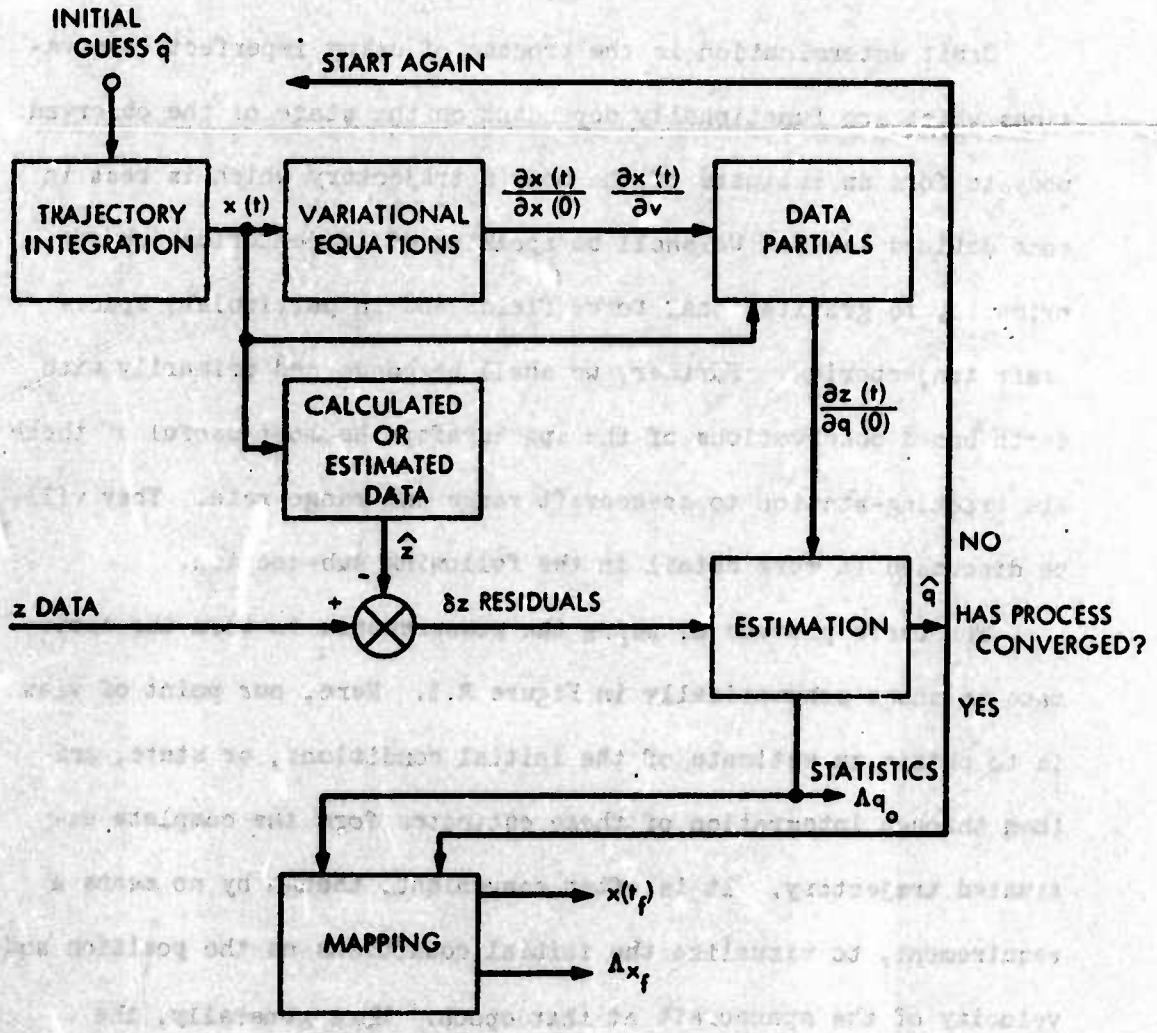


Figure A.1. Functional Diagram of Orbit Determination Process.

where x(0) is the spacecraft state (position and velocity) a six vector

v are parameters which influence the propogation of spacecraft state (e.g. mass of the earth, solar pressure)

s are parameters which affect only the data (e.g. station locations, data biases)

It is important to restrict the members of q to only those parameters where the uncertainty in their true values significantly affect the problem at hand. Parameters that are relatively well known or whose involvement in the problem is minimal are treated as perfectly known and do not appear in q (e.g. mass of Pluto for a Mars Orbiter problem).

To start the process, an initial guess of q must be furnished. This is used to integrate the non-linear equations of motion and produce a nominal probe ephemeris. Then, a linearization about this nominal takes place to calculate the effects of small variations in q. That is, the system of equations

$$\dot{\phi}(t,t_o) = A\phi(t,t_o)$$

where

$$A = \begin{bmatrix} \frac{\partial \dot{x}}{\partial x} & \frac{\partial \dot{x}}{\partial v} \\ \\ \frac{\partial \dot{v}}{\partial x} & \frac{\partial \dot{v}}{\partial v} \end{bmatrix}$$ is evaluated about the nominal

in a straightforward manner. Since the v's are formulated as constant parameters of the system, the lower partitions of A are zero and the ϕ obtained has the special form .

$$\Phi(t,t_o) = \left[\begin{array}{c|c} U & V \\ \hline 0 & I \end{array}\right]$$

thus

$$\delta x(t) = \left[\begin{array}{c|c} U & V \end{array}\right] \begin{bmatrix} \delta x(0) \\ \delta v \end{bmatrix}$$

and

$$U = \frac{\partial x(t)}{\partial x(0)} \quad ; \quad V = \frac{\partial x(t)}{\partial v}$$

The probe trajectory is also used to calculate "estimated" observables. This is done by simply asking the physically meaningful question, "If the actual trajectory were exactly as I have guessed it to be, and a perfect observation were taken, what would be the value of that observation?" The estimated observation is differenced with that actually received, producing residuals that are fed directly into the estimation process.

In addition to the residuals, we have need of the partial derivatives of the data with respect to the total parameter vector, q. Both the U and V matrices already defined are needed in this final computation. Let the data received at time t be

$$z(t) \triangleq h(t, x(t; x(0), v), v, \varepsilon)$$

then

$$\frac{\partial z(t)}{\partial x(0)} = \frac{\partial h}{\partial x(t)} \frac{\partial x(t)}{\partial x(0)} = U^T \frac{\partial h}{\partial x(t)}$$

$$\frac{\partial z(t)}{\partial v} = \frac{\partial h}{\partial x(t)} \frac{\partial x(t)}{\partial v} + \frac{\partial h}{\partial v} \Bigg|_{x(t) \text{ const}}$$

$$= V^T \frac{\partial h}{\partial x(t)} + \frac{\partial h}{\partial v}$$

and

$$\frac{\partial z(t)}{\partial s} = \frac{\partial h}{\partial s}$$

The v enters both explicitly in h and as an argument of x(t) since certain parameters can affect both the propogation of the spacecraft state and the data directly. These are unusual however and do not arise often; perhaps the best example is the ephemeris of the Earth. Its position obviously affects the trajectory, but in addition, a change in the Earth's position moves the tracking station and contributes directly to a change in the data.

Together, the partials and the residuals are used in the estimation process to form the estimate of q, $\hat{q}$. Along with this estimate will also be statistics on the errors in the estimate in the form of a covariance matrix, Λ_q, for the second moments (the first moments, generally, are zero). Usually, the quantity of prime interest is not the initial conditions, but the spacecraft trajectory near or at a distant target body. If the later is the case, the estimated $\hat{q}$ is used to produce a new complete trajectory. In addition, the statistics on the state can also be propogated by

$$\Lambda_x^{(t)} = \left[\begin{array}{c|c} U & V \end{array}\right] \Lambda_q \left[\begin{array}{c|c} U & V \end{array}\right]^T$$

where U and V are understood to have arguments t, t_o.

II. Earth-Based Radio Tracking Data

Definitions and Characteristics

We include here a brief description of the physical system employed by the Deep Space Network (DSN) of the Jet Propulsion Laboratory and a measure of the inherent accuracy of that system.

Figure A.2 shows the basic building blocks for a coherent range and doppler system. This figure could have been drawn as separate diagrams and, indeed, separate systems could have been mechanized for each of the data types. It is useful, however, to emphasize not only the common elements of the hardware system, but the conceptual relationship between the two data types. Basically, the range system measures the roundtrip light time between the probe and station, while the doppler system measures the time rate of change of this quantity. The latter is accomplished by sending a signal having a stable frequency to the spacecraft. The signal received at the spacecraft is re-broadcast to the station and is received with a frequency shift proportional to the relative range-rate. The doppler tone produced is simply integrated by means of an electronic counter which results in the integral of the range-rate.

At the same time, the carrier is phase modulated with the ranging code. The code broadcast can have a variety of mathematical structures depending on the formulation used. A particularly interesting code, however, one currently being used by the DSN, is the so-called multi-component "pseudorandom code." Although this particluar code is generated deterministically and has a finite length

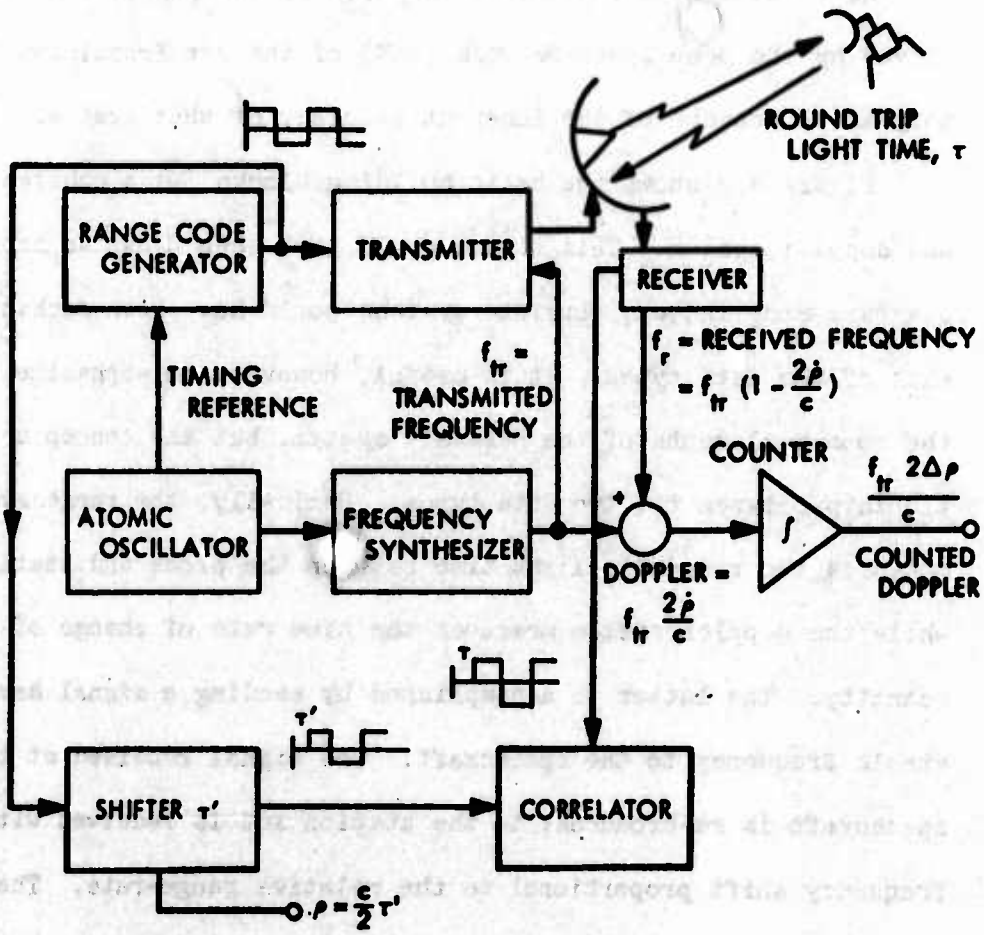


Figure A.2. Coherent Range, Doppler Ground Tracking System.

before repetition, it takes on many of the characteristics of a purely random sequence of 1's and 0's. When this signal is returned to the station, it has shifted in time corresponding to the instantaneous round-trip light time involved. The signal is then auto-correlated with the code being generated. The generated code is also shifted in time as shown in Figure A.2. Here, the properties of the pseudo-random sequence are used, and the correlation computed will be uniformly low, unless the generated code is shifted in time by the same amount that the received code is shifted by its round-trip transit time. By carefully constructing the code from subcodes of gradually increasing lengths, this search for maximum correlation can be done efficiently and need not require a trial correlation with every possible shift time. The amount of shift required to produce this maximum correlation becomes the raw data type proportional to the spacecraft range. Therefore, the counted doppler is seen as the integral of the range-rate while the ranging system supplies the arbitrary constant of integration. These are the raw data types actually used.

This system, particularly the doppler portion, is relatively unsusceptible to instrument error. The resulting high inherent data accuracy has deferred the need for on-board tracking systems, and has propelled the Earth-based system as the prime means of orbit determination far beyond the limits originally envisioned when planetary exploration was first undertaken.

In order to establish a means of comparison, consider tracking a spacecraft having a typical Earth-relative velocity of 10 km/sec. Since the raw data type is range change, and this is measured

continuously over periods up to 12 hr (a single pass), the measurement will be on the order of 500,000 km.

Perhaps the most visible noises impressed on the data are receiver jitter and rather high-frequency variations arising from the master oscillator instability. These errors are visible because of the high-frequency content, and will be in evidence after even a rudimentary smoothing of the data. Their excursions are small, however, not amounting to more than a few centimeters, or an upper bound of one part in 10^9 of the basic measurement. In addition, because the noise is high frequency tends to make it easily filtered out in any orbit determination process, and, therefore, it does not seriously affect the estimates.

Low-frequency errors, from a range-rate bias to frequencies approximately 10 times the frequency of the Earth's rotation, are of greater importance. Fortunately, the instrument is not particularly susceptible to these errors. Consider again Figure A.2 and note that a bias in the oscillator frequency will cause a shift both in the transmitted and received frequencies, quantities that are differenced to produce the doppler tone. The effect of the bias is nearly cancelled by this operation. If, on the other hand, a change in the frequency occurs, the doppler will be affected directly by the magnitude of that change. The shift in the doppler tone is integrated to produce a range-change error. A cancellation occurs later, however, when the frequency change reappears in the received signal; this time it enters the data in the opposite sense. Therefore, an oscillator bias is of no consequence; it is the variations in the transmitted

frequency that produce first-order effects. Current rubidium standard performance of drifts up to 5 parts in 10^{12} over a pass will produce errors of 80 cm for a 20-minute round-trip light time (as shown in Chapter II).

The approximate doppler equations shown in Figure A.2 give the impression that the uncertainty in the speed of light, c, would give rise to a bias in the range-rate measurement. This arises, however, only because c enters from a desire to express velocities in metric units; it is not a requirement. All measurements and ephemerides are fundamentally taken and calculated in light-seconds, the distance light travels in unit time. Therefore, c is a defined constant for convenience only; it could as well be unity.

There are a variety of sources giving rise to errors in the remainder of the low-frequency region. In total, they add to a standard range-change error of approximately 1.5 m during a single doppler tracking pass, or using the original yardstick of 500,000 km, 3 parts in 10^9.

The current planetary ranging system is capable of measuring the station-spacecraft distance to within a standard error of approximately 15 m bias plus a 5 m random component. These numbers justify the earlier statement that the ranging data supply the constant of range-rate integration; i.e., on the basis of a single pass, the doppler gives a considerably more accurate measurement of range change than does the difference of two range measurements. Taking a longer view, the situation is reversed since doppler errors willcontinue to accumulate but the difference between two range measurements even after

several days will still be accurate, to, say, $\sqrt{2} \times 5$ m. To capitalize on this feature, one must pay careful attention to maintaining the calibration of the ranging system to ensure that the bias is literally that, and not a slowly varying quantity.

III. The Estimation Process

We now turn our attention to the estimation process itself. Let the data be represented as a linear function of the state (actually, of course, the perturbation from nominal in the state) plus additive noise, i.e.

$$z = Hx + \epsilon \tag{A.1}$$

where z is an m vector

x is the (unknown) state, an n vector*

and H is an mxn matrix of partial derivatives

ϵ is an m vector of the measurement error.

It is assumed that $E[\epsilon] = 0$ and $E[\epsilon\,\epsilon^T] = \Lambda_\epsilon$ where E is the expected value operator; superscript T denotes the transpose. Our objective is to find a linear, un-biased minimum variance estimate of x, $\hat{x}$. Or stated precisely

i. $\hat{x} = Bz$ (linearity)

ii. $E[\hat{x}] = x$ (unbiased)

iii. $\Lambda_x = E[(x-\hat{x})(x-\hat{x})^T]$ is a minimum

Substituting (A.1) into i and employing ii

$$\hat{x} = B(Hx + \epsilon) \tag{A.2}$$

$$E[\hat{x}] = BHx = x \text{ if and only if}$$

$$BH = I \tag{A.3}$$

---

*It is convenient to now drop the distinction between x, v, and s and treat all parameters in a unified fashion. Thus, x here is the same as δq in Section I and is used to conform with more universal notation.

Substituting (A.2) and (A.3) into iii gives

$$\Lambda_x = E\left[B\epsilon\ \epsilon^T B^T\right] = B\Lambda_\epsilon B^T \tag{A.4}$$

Equation (A.4) must be minimized with respect to B but observing the constraint, (A.3). Thus, adjoining this constraint with a matrix Lagrange multiplier to (A.4)

$$\Lambda_x = B\Lambda_\epsilon B^T + (BH-I)\lambda + \lambda^T(H^T B^T - I) \tag{A.5}$$

In order for (4) to be a minimum, the variation of (5) must be zero.

$$\delta\Lambda_x = \delta B(\Lambda_\epsilon B^T + H\lambda) + (B\Lambda_\epsilon + \lambda^T H^T)\delta B^T = 0 \tag{A.6}$$

Evidently

$$B\Lambda_\epsilon + \lambda^T H^T = 0$$

or

$$B = -\lambda^T H^T \Lambda_\epsilon^{-1}$$

and since $BH = I$

$$\lambda^T = \lambda = -(H^T \Lambda_\epsilon^{-1} H)^{-1}$$

thus

$$\hat{x} = (H^T \Lambda_\epsilon^{-1} H)^{-1} H^T \Lambda_\epsilon^{-1} z \tag{A.7}$$

and

$$\Lambda_x = (H^T \Lambda_\epsilon^{-1} H)^{-1} \tag{A.8}$$

These are the familiar weighted least-squares or batch minimum variance estimation formulae. They have been in use since Gauss with very little modification and are still the mainstay of many sophisticated orbit determination problems in use throughout the country.

With very little additional work we can introduce the concept of apriori information on the state prior to data taking.

Suppose we have an initial guess as to the value of x, in the form

$$\bar{x} = x + \epsilon_x$$

where ϵ_x is mean zero with covariance Λ_{ap}.

This can be considered as simply additional data related to x through an identity transformation. Since usually (though not necessarily) the initial guess at the trajectory will correspond to the apriori data, in our notation that data becomes zero; i.e.

$$\begin{bmatrix} 0 \\ z \end{bmatrix} = \begin{bmatrix} I \\ H \end{bmatrix} x + \begin{bmatrix} \epsilon_x \\ \epsilon \end{bmatrix} \tag{A.9}$$

Substituting (A.9) into (A.7) and (A.8) yields

$$\hat{x} = \left\{ \begin{bmatrix} I & H^T \end{bmatrix} \begin{bmatrix} \Lambda_{ap}^{-1} & 0 \\ 0 & \Lambda_\epsilon^{-1} \end{bmatrix} \begin{bmatrix} I \\ H \end{bmatrix} \right\}^{-1} \begin{bmatrix} I & H^T \end{bmatrix} \begin{bmatrix} \Lambda_{ap}^{-1} & 0 \\ 0 & \Lambda_\epsilon^{-1} \end{bmatrix} \begin{bmatrix} 0 \\ z \end{bmatrix}$$

$$\hat{x} = (\Lambda_{ap}^{-1} + H^T \Lambda_\epsilon^{-1} H)^{-1} H^T \Lambda_\epsilon^{-1} z \tag{A.7}'$$

and

$$\Lambda_x = (\Lambda_{ap}^{-1} + H^T \Lambda_\epsilon^{-1} H)^{-1} \tag{A.8}'$$

Sequential or Kalman-Bucy Filtering

A version of the Gauss Markov Theorem (Reference 24) states that if two mean zero random variables, x and z, have covariances Λ_{ap} and Λ_z respectively and cross covariance Λ_{xz}, then the minimum variance, linear, unbiased estimate of x given z is:

$$\hat{x} = \Lambda_{xz} \Lambda_{zz}^{-1} z \tag{A.10}$$

and the resulting covariance in the error $\hat{x}$ is given by

$$\Lambda_x = \Lambda_{ap} - \Lambda_{xz}\Lambda_{zz}^{-1}\Lambda_{xz}^T \qquad (A.11)$$

If z is again considered the data in accordance with (1), and x has prior uncertainty Λ_{ap}, then

$$\Lambda_{xz} = \Lambda_{ap}H^T$$

$$\Lambda_{zz} = (H\,\Lambda_{ap}H^T + \Lambda_\epsilon)$$

Thus (A.10) and (A.11) become:

$$\hat{x} = \Lambda_{ap}H^T(H\,\Lambda_{ap}H^T + \Lambda_\epsilon)^{-1}z \qquad (A.10)'$$

$$\Lambda_x = \Lambda_{ap} - \Lambda_{ap}H^T(H\,\Lambda_{ap}H^T + \Lambda_\epsilon)^{-1}H\,\Lambda_{ap} \qquad (A.11)'$$

This in essence, is the Kalman-Bucy filter in discrete form for the very simplest of circumstances (Reference 25). We do not present in detail here the equations for the full recursive algorithm including mapping between data times because they are lengthy and only serve to obscure the elegance of the principal results - all of the basic ideas are contained in these two equations. In words, the procedure becomes:

If (A.10)' and (A.11)' represent the results from the first data point (or points) and it is desired to process more data -

1) replace Λ_{ap} with Λ_x; corrupt Λ_x by the addition of a positive definite, symmetric matrix if random process noise is entering the system
2) subtract from the new data $H'\hat{x}$, which is the expected value of the new data (H' is the new partial derivative matrix)
3) re-execute (10)', (11)'; the resulting $\hat{x}$ is the amount the

estimate should be changed from the previous value.

Both the least-squares and sequential methods are unbiased, minimum variance techniques. If this is so, it must be possible to show the equivalence between (A.7)', (A.8)', and (A.10)', (A.11)'. We begin with an apparent diversion. Suppose we have two matrices, partitioned commensurably, one the inverse of the other

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} = \left[\begin{array}{c|c} I & O \\ \hline O & I \end{array}\right]$$

Then

$$A\alpha + B\gamma = I \qquad \text{(i)}$$

$$C\alpha + D\gamma = O \qquad \text{(ii)}$$

First, solving (ii) for γ, and using in (i), we obtain

$$\gamma = -D^{-1}C\alpha$$

$$\alpha = (A - BD^{-1}C)^{-1} \qquad \text{(iii)}$$

$$\gamma = D^{-1}C(A - BD^{-1}C)^{-1} \qquad \text{(iv)}$$

Second, solving (i) for α and using in (ii)

$$\alpha = A^{-1} - A^{-1}B\gamma$$

$$\alpha = A^{-1} - A^{-1}B(CA^{-1}B - D)^{-1}CA^{-1} \qquad \text{(v)}$$

$$\gamma = (CA^{-1}B-D)^{-1}CA^{-1} \qquad \text{(vi)}$$

Equating (iii), (v) and (iv), (vi) we have finally

$$(A - BD^{-1}C)^{-1} = A^{-1} - A^{-1}B(CA^{-1}B-D)^{-1}CA^{-1} \qquad \text{(A.12)}$$

$$(CA^{-1}B-D)^{-1}CA^{-1} = -D^{-1}C(A - BD^{-1}C)^{-1} \qquad \text{(A.13)}$$

Now for (A.12) if we let

$$D \rightarrow -\Lambda_{ap}^{-1}$$

$$C \rightarrow H^T$$

$$B \rightarrow H$$

$$A \rightarrow \Lambda_\epsilon$$

Then

$$(H^T\Lambda_\epsilon^{-1}H + \Lambda_{ap}^{-1})^{-1}H^T\Lambda_\epsilon^{-1} = \Lambda_{ap}H^T(\Lambda_\epsilon + H\Lambda_{ap}H^T)^{-1}$$

showing the equivalence of the two filter equations.

Finally if for (A.13) we let

$$A \rightarrow \Lambda_{ap}^{-1}$$

$$D^{-1} \rightarrow -\Lambda_\epsilon^{-1}$$

$$B \rightarrow H^T$$

$$C \rightarrow H$$

Then

$$(\Lambda_{ap}^{-1} + H^T\Lambda_\epsilon^{-1}H)^{-1} = \Lambda_{ap} - \Lambda_{ap}H^T(H\Lambda H^T + \Lambda_\epsilon)^{-1}H\Lambda_{ap}$$

showing the equivalence of the two covariance equations.

Square-Root Filtering

The basic motivation of square-root filtering, which can be used either with the least-squares or sequential techniques is that by operating with either the square root of the covariance or information (inverse covariance) matrices, numerical integrity is improved. This is best illustrated by a specific example. Suppose, we wished to solve the following equation for x

$$z = Hx$$

where

$$H = \begin{bmatrix} 1 & 1 \\ 1 & 1+\epsilon \end{bmatrix}$$

on a digital computer in single precision (8 decimal places). To do this in the usual way would involve taking the determinant,

$$\det H = 1 + \epsilon - 1$$

If ϵ were, say $\approx 10^{-6}$, the calculated determinant would have two place accuracy and a reasonable answer would result. If, however, we were to solve this problem in a least-square sense which theoretically yields the same answer, i.e., if we were to compute

$$x = (H^T H)^{-1} H^T z,$$

we would have to essentially find the determinant of $H^T H$, which is readily seen to be

$$\det H^T H = 4 + 4\epsilon + 2\epsilon^2 - 4 - 4\epsilon - \epsilon^2$$

If ϵ were $\approx 10^{-6}$, this would compute as zero, and the matrix could not be inverted. Since for accurate orbit determination it is necessary to include as estimated variables all parameters that contribute information to the data, the problem of a nearly rank deficient H is chronic. Thus the motivation is high to avoid the $H^T H$ operation and to find ways of dealing with H directly even (particularly) when it is non-square. We present the outline of one such technique here that employs a generalization of the Householder transformations originally used for inverting square matrices by first triangularizing and then inverting by backwards substitution.*

We begin with the least-squares data model

$$z = Hx + \epsilon \qquad (A.14)$$

*Definitions: By the square-root of a square matrix Q, $Q^{\frac{1}{2}}$, we simply mean the (non-unique) matrix such that $Q^{\frac{1}{2}} Q^{\frac{1}{2}T} = Q$. An upper (lower) triangular matrix has all zero elements below(above) the main diagonal.

where the same definitions apply as before.

If (A.14) is pre-multiplied by $\Lambda_\epsilon^{-\frac{1}{2}}$ and

$$n \triangleq \Lambda_\epsilon^{-\frac{1}{2}} \epsilon$$

$$\Lambda_\epsilon^{-\frac{1}{2}} z = \Lambda_\epsilon^{-\frac{1}{2}} x + n \qquad \text{(A.15)}$$

where now $E[n] = 0$, $E[n\, n^T] = I$

We seek an estimate of the state x and the noise, n, such that $\hat{n}$ is minimum subject to Equation (A.15), i.e.

$$\Lambda_\epsilon^{-\frac{1}{2}} H \hat{x} + \hat{n} = \Lambda_\epsilon^{-\frac{1}{2}} z \qquad \text{(A.16)}$$

It is possible, through an orthonormal transformation P to find a new basis such that the column vectors of $\Lambda_\epsilon^{-\frac{1}{2}} H$ are contained within the sub-space formed by the first n vectors of the new basis, i.e.

$$P\Lambda_\epsilon^{-\frac{1}{2}} H = \begin{matrix} n \\ m-n \end{matrix} \begin{bmatrix} R \\ -- \\ 0 \end{bmatrix}$$

Equation (A.16) becomes

$$\begin{bmatrix} R \\ \text{----} \\ 0 \end{bmatrix} \hat{x} + \begin{bmatrix} \hat{n}_x \\ \text{---} \\ \hat{n}_z \end{bmatrix} = \begin{bmatrix} z_1' \\ \text{----} \\ z_2' \end{bmatrix} \qquad \text{(A.17)*}$$

Since P is orthonormal, $\left\| \begin{bmatrix} \hat{n}_x \\ \text{---} \\ \hat{n}_z \end{bmatrix} \right\| = \|\hat{n}\|$ and the problem is still to satisfy (A.17) and minimize the $\left\| \begin{bmatrix} \hat{n}_x \\ \text{---} \\ \hat{n}_z \end{bmatrix} \right\|$. This solution is simply

---

*The conditions on P thus far imposed make P non-unique. We can facilitate the later inversion of R and make P unique if, in addition we require that R be upper triangular

$$\hat{h}_x = 0$$

$$\hat{h}_z = z_2'$$

$$\hat{x} = R^{-1} z_1'$$

To show equivalence with the least square formulation just discussed we observe:

1. The solution is linear
2. it is unbiased, since

$$\hat{x} = R^{-1} z_1' = R^{-1} R x + R^{-1} P n$$

$$E[\hat{x}] = Ix + R^{-1} P E[n] = x$$

3. the covariance of the estimate is still the same.

$$x - \hat{x} = R^{-1} P n$$

$$E[(x-\hat{x})(x-\hat{x})^T] \triangleq \Lambda_x = R^{-1} P E[n\, n^T] P^T R^{-1^T} = R^{-1} R^{-1^T}$$

since $P\, P^T = I$ from the consequence of the orthonormality.

But

$$\left[\begin{array}{c|c} R^T & 0 \end{array}\right] \left[\begin{array}{c} R \\ \hline 0 \end{array}\right] = H^T \Lambda_e^{-\frac{1}{2}} P\, P^T \Lambda_e^{-\frac{1}{2}} H$$

or

$$R^{-1} R^{-1^T} = (H^T \Lambda_e^{-1} H)^{-1}$$

which is the same as the weighted least square covariance and, by the proof of the previous sub-section, as the sequential formula.

BIBLIOGRAPHY

1. Curkendall, D. W., and others, "Planetary Navigation - the New Challenges," Astronautics & Aeronautics, 8: 26-70, May, 1970.

2. Hamilton, T. W., and W. G. Melbourne, "Information Content of a Single Pass of Doppler Data from a Distant Spacecraft," Jet Propulsion Laboratory, Space Programs Summary Number 37-39, May 31, 1966, Vol. III.

3. Papoulis, Athanasios, Probability, Random Variable and Stochastic Processes, New York, McGraw-Hill, 1965.

4. Curkendall, D. W., "Orbit Accuracy as a Function of Doppler Sample Rate for Several Data Taking and Processing Modes," Jet Propulsion Laboratory, Space Programs Summary Number 37-38, March 31, 1966, Vol. III, p. 20-24.

5. Light, J. O., "An Investigation of the Orbit Redetermination Process Following the First Midcourse Maneuver," Jet Propulsion Laboratory, Space Programs Summary Number 37-33, June 30, 1965, Vol. IV.

6. Curkendall, D. W., "Six Parameter Model Characterizing Doppler Information Received While Tracking a Spacecraft," Jet Propulsion Laboratory, Space Programs Summary Number 37-47, September 30, 1967.

7. Curkendall, D. W., and S. R. McReynolds,"A Simplified Approach for Determining the Information Content of Radio Tracking Data," Journal of Spacecraft & Rockets, 6:520-525, May, 1969.

8. Ondrasik, V. J., "A First Order Theory for Use In Investigating the Information Content Contained in a Few Days of Radio Tracking Data," Jet Propulsion Laboratory, Technical Report to be published.

9. Householder, A. S., "A Unitary Triangularization of a Nonsymmetric Matrix," Journal of Association of Computing Machinery, 5: 339-342, 1958.

10. Businger, P. and G. H. Golub, "Linear Least Squares Solutions by Householder Transformations," Numer. Math., 7: 269-276, 1965.

11. Hansen, R. J., and C. L. Lawson, "Extensions and Applications of the Householder Algorithm for Solving Linear Least Squares Problems," Mathematics of Computation, 23: 787-812, October, 1969.

12. Dyer, P., and S. R. McReynolds, "The Extension of Square Root Filtering to Include Process Noise," Journal of Optimization: Theory & Applications, 3: 92-105, June, 1969.

13. Dyer, P., "Formulas for Implementation of the Householder Algorithm in the DPODP," Jet Propulsion Laboratory, Space Programs Summary Number 37-58, May, 1969, Vol. II.

14. Soong, T. T., "On A Priori Statistics in Minimum Variance Estimation Problems," Journal of Basic Engineering, 87D: 109-112, March, 1965.

15. Nishimura, T., "On the A Priori Information In Sequential Estimation Problems," IEEE Transactions on Automatic Control, AC-11: 197-204, April, 1966.

16. Nishimura, T., "Modeling Errors in Kalman Filters," Theory and Applications of Kalman Filtering, edited by C. T. Leondes, NATO, AGARDograph No. 139, February, 1970.

17. Gordon, H. J., and others, "The Mariner VI and VII Flight Paths and Their Determination from Tracking Data," Jet Propulsion Laboratory, Technical Memorandum Number 33-469, December 1, 1970.

18. Jordan, J. F., G. A. Madrid, and G. E. Pease, "The Effects of Major Error Sources on Planetary Spacecraft Navigation Accuracies," Paper presented at the AAS/AIAA Astrodynamics Conference, August 19-21, 1970, Santa Barbara, California.

19. Warner, M. R., and others, "The Double Precision Orbit Determination Program," Jet Propulsion Laboratory Publication Series Number 900-224; Rev. A., November, 1966, Vol. I-VIII.

20. Plamondon, J. A., "The Measurement of the Solar Constant with Mariner VI and VII," Jet Propulsion Laboratory, Space Programs Summary Number 37-59, July, 1969, Vol. III.

21. Pease, G. E., and others, "The Mariner V Flight Path and Its Determination from Tracking Data," Jet Propulsion Laboratory, Technical Report Number 32-1362, July, 1969.

22. Anderson, J. D., L. Efron, and S. K. Wong, "Martian Mass and Earth-Moon Mass Ratio from Coherent S-Band Tracking of Mariners 6 and 7," Science, 167: 277-279, January 16, 1970.

23. Jordan, J. F., "Orbit Determination for Powered Flight Space Vehicles on Deep Space Missions," Journal of Spacecraft & Rockets, 6: 545-550, May, 1969.

24. Scheffe, Henry, The Analysis of Variance, New York, John Wiley & Sons, 1959.

25. Kalman, R. E., and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," Journal of Basic Engineering Transactions of the ASME, 83: 95-108, March, 1961.

26. Gordon, H. J., S. K. Wong, and V. J. Ondrasik, "Analysis of Mariner VII Preencounter Anomaly," Paper presented at the AAS/AIAA Astrodynamics Conference, August 19-21, 1970, Santa Barbara, California.